# OIV-MA-AS1-12 Practical guide for the validation, quality control, and uncertainty assessment of an alternative oenological analysis method

Contents

**1. Purpose**
**2. Preamble and scope**
**3. General vocabulary**
**4. General principles**
4.1 Methodology
4.2 Definition of measurement error
**5. Validating a method**
5.1 Methodology
5.2 Section one: Scope of method
5.2.1 Definition of analyzable matrices
5.2.2 Detection and quantification limit
5.2.2.1 Normative definition
5.2.2.2 Reference documents
5.2.2.3 Application
5.2.2.4 Procedure
5.2.2.4.1 Determination on blank
5.2.2.4.1.1 Scope
5.2.2.4.1.2 Basic protocol and calculations
5.2.2.4.2 Approach by linearity study
5.2.2.4.2.1 Scope
5.2.2.4.2.2 Basic protocol and calculations
5.2.2.4.3 Graphic approach based on the background noise of the recording
5.2.2.4.3.1 Scope
5.2.2.4.3.2 Basic protocol and calculation
5.2.2.4.4 Checking a predetermined quantification limit
5.2.2.4.4.1 Scope
5.2.2.4.4.2 Basic protocol and calculation
5.2.3 Robustness

5.2.3.1 Definition
5.2.3.2 Determination
5.3 Section two: systematic error study
5.3.1 Linearity study
5.3.1.1 Normative definition
5.3.1.2 Reference documents
5.3.1.3 Application
5.3.1.4 ISO 11095-type approach
5.3.1.4.1 Basic protocol
5.3.1.4.2 Calculations and results
5.3.1.4.2.1 Defining the regression model
5.3.1.4.2.2 Estimating parameters
5.3.1.4.2.3 Charts
5.3.1.4.2.4 Test of the linearity assumption
5.3.1.4.2.4.1 Definitions of errors linked to calibration
5.3.1.4.2.4.2 Fischer-Snedecor test
5.3.1.5 ISO 8466-type approach
5.3.1.5.1 Basic protocol
5.3.1.5.2 Calculations and results
5.3.1.5.2.1 Defining the linear regression model
5.3.1.5.2.2 Defining the polynomial regression model
5.3.1.5.2.3 Comparing residual standard deviations
5.3.2 Specificity
5.3.2.1 Normative definition
5.3.2.2 Application
5.3.2.3 Procedures
5.3.2.3.1 Standard addition test
5.3.2.3.1.1 Scope
5.3.2.3.1.2 Basic protocol
5.3.2.3.1.3 Calculations and results
5.3.2.3.1.3.1 Study of the regression line **r = a + b.v**
5.3.2.3.1.3.2 Analysis of the results
5.3.2.3.1.3.3 Overlap line graphics

5.3.2.3.2 Study of the influence of other compounds on the measurement result
5.3.2.3.2.1 Scope
5.3.2.3.2.2 Basic protocol and calculations
5.3.2.3.2.3 Interpretation
5.3.3 Study of method accuracy
5.3.3.1 Presentation of the step
5.3.3.1.1 Definition
5.3.3.1.2 General principles
5.3.3.1.3 Reference documents
5.3.3.2 Comparison of the alternative method with the OIV reference method
5.3.3.2.1 Scope
5.3.3.2.2 Accuracy of the alternative method compared with the reference method
5.3.3.2.2.1 Definition
5.3.3.2.2.2 Scope
5.3.3.2.2.3 Basic protocol and calculations
5.3.3.2.2.4 Interpretation
5.3.3.3 Comparison by interlaboratory tests
5.3.3.3.1 Scope
5.3.3.3.2 Basic protocol and calculations
5.3.3.3.3 Interpretation
5.3.3.4 Comparison with reference materials
5.3.3.4.1 Scope
5.3.3.4.2 Basic protocol and calculations
5.3.3.4.3 Interpretation
5.4 Section three: random error study
5.4.1 General principle
5.4.2 Reference documents
5.4.3 Precision of the method
5.4.3.1 Definition
5.4.3.2 Scope
5.4.3.3 General theoretical case
5.4.3.3.1 Basic protocol and calculations
5.4.3.3.1.1 Calculations with several test materials

5.4.3.3.1.2 Calculations with 1 test material
5.4.3.4 Repeatability
5.4.3.4.1 Definitions
5.4.3.4.2 Scope
5.4.3.4.3 Basic protocol and calculations
5.4.3.4.3.1 General case
5.4.3.4.3.2 Particular case applicable to only 1 repetition
5.4.3.4.4 Comparison of repeatability
5.4.3.4.4.1 Determination of the repeatability of each method
5.4.3.4.4.2 Fischer-Snedecor test
5.4.3.5 Intralaboratory reproducibility
5.4.3.5.1 Definition
5.4.3.5.2 Scope
5.4.3.5.3 Basic protocol and calculations
**6. Quality control of analysis methods (IQC)**
6.1 Reference documents
6.2 General principles
6.3 Reference materials
6.4 Checking the analytical series
6.4.1 Definition
6.4.2 Checking accuracy using reference materials
6.4.3 Intraseries precision
6.4.4 Internal standard
6.5 Checking the analysis system
6.5.1 Definition
6.5.2 Shewhart chart
6.5.2.1 Data acquisition
6.5.2.2 Presentation of results and definition of limits
6.5.2.3 Using the Shewhart chart
6.5.3 Internal comparison of analysis systems
6.5.4 External comparison of the analysis system
6.5.4.1 Analysis chain of interlaboratory comparisons
6.5.4.2 Comparison with external reference materials

6.5.4.2.1 Standard uncertainty of reference material
6.5.4.2.2 Defining the validity limits of measuring reference material
**7. Assessment of measurement uncertainty**
7.1 Definition
7.2 Reference documents
7.3 Scope
7.4 Methodology
7.4.1 Definition of the measurand, and description of the quantitative analysis method
7.4.2 Critical analysis of the measurement process
7.4.3 Estimation calculations of standard uncertainty (intralaboratory approach)
7.4.3.1 Principle
7.4.3.2 Calculating the standard deviation of intralaboratory reproducibility
7.4.3.3 Estimating typical sources of systematic errors not taken into account underμ reproducibility conditions
7.4.3.3.1 Gauging error (or calibration error)
7.4.3.3.1.1 Procedure
7.4.3.3.1.2 Calculations and results
7.4.3.3.1.3 Estimating the standard uncertainty associated the gauging line (or calibration line)
7.4.3.3.2 Bias error
7.4.3.3.2.1 Methods adjusted with only one certified reference material
7.4.3.3.2.2 Methods adjusted with several reference materials (gauging ranges etc)
7.4.3.3.3 Matrix effect
7.4.3.3.3.1 Definition
7.4.3.3.4 Sample effect
7.4.4 Estimating standard uncertainty by interlaboratory tests
7.4.4.1 Principle
7.4.4.2 Using the standard deviation of interlaboratory and intramethod reproducibility $SR_{inter}$ (method)
7.4.4.3 Using the standard deviation of interlaboratory and intermethod reproducibility $SR_{inter}$

7.4.4.4 Other components in the uncertainty budget
7.5 Expressing expanded uncertainty

## 1. Purpose

The purpose of this guide is to assist oenological laboratories carrying out serial analysis as part of their validation, internal quality control and uncertainty assessment initiatives concerning the standard methods they use.

## 2. Preamble and scope

International standard ISO 17025, defining the "General Requirements for the Competence of Testing and Calibration Laboratories", states that the accredited laboratories must, when implementing an alternative analytical method, make sure of the quality of the results obtained. To do so, it indicates several steps. The first step consists in defining the customers' requirements concerning the parameter in question, in order to determine, thereafter, whether the method used meets those requirements. The second step includes initial validation for non-standardized, modified or laboratory-developed methods. Once the method is applied, the laboratories must use inspection and traceability methods in order to monitor the quality of the results obtained. Finally, they must assess the uncertainty of the results obtained.

In order to meet these requirements, the laboratories have a significant reference system at their disposal comprising a large number of international guides and standards. However, in practice, the application of these texts is delicate since, because they address every category of calibration and test laboratory, they remain very general and presuppose, on behalf of the reader, in-depth knowledge of the mathematical rules applicable to statistical data processing.

This guide is based on this international reference system, taking into account the specific characteristics of oenology laboratories routinely carrying out analyses on series of must or wine samples. Defining the scope of application in this way enabled a relevant choice of suitable tools to be made, in order to retain only those methods most suitable for that scope. Since it is based on the international reference system, this guide is therefore strictly compliant with it. Readers, however, wishing to study certain points of the guide in greater detail can do so by referring to the international standards and guides, the references for which are given in each chapter.

The authors have chosen to combine the various tools meeting the requirements of the ISO 17025 standard since there is an obvious solution of continuity in their application, and the data obtained with certain tools can often be used with the

others. In addition, the mathematical resources used are often similar.

The various chapters include application examples, taken from oenology laboratories using these tools.

It is important to point out that that this guide does not pretend to be exhaustive. It is only designed to present, in as clear and applicable a way as possible, the contents of the requirements of the ISO 17025 standard and the basic resources that can be implemented in a routine laboratory to meet them. Each laboratory remains perfectly free to supplement these tools or to replace them by others that they consider to be more efficient or more suitable.

Finally, the reader's attention should be drawn to the fact that the tools presented do not constitute an end in themselves and that their use, as well as the interpretation of the results to which they lead, must always be subject to critical analysis. It is only under these conditions that their relevance can be guaranteed, and laboratories will be able to use them as tools to improve the quality of the analyses they carry out.

## 3. General vocabulary

The definitions indicated below used in this document result from the normative references given in the bibliography.

**Analyte**

Object of the analysis method

**Blank**

Test carried out in the absence of a matrix (reagent blank) or on a matrix which does not contain the analyte (matrix blank).

**Bias**

Difference between the expected test results and an accepted reference value.

**Uncertainty budget**

The list of uncertainty sources and their associated standard uncertainties, established in order to assess the compound standard uncertainty associated with a measurement result.

**Gauging (of a measuring instrument)**

Material positioning of each reference mark (or certain principal reference marks only) of a measuring instrument according to the corresponding value of the measurand.

NOTE "gauging" and "calibration" are not be confused

**Repeatability conditions**

Conditions where independent test results are obtained with the same method on

identical test items in the same laboratory by the same operator using the same equipment within short intervals of time.

**Reproducibility conditions (intralaboratory)**

Conditions where independent test results are obtained with the same method on identical test items in the same laboratory by the same or different operator(s) using different gauges on different days.

**Experimental standard deviation**

For a series of ***n*** measurements of the same measurand, the quantity ***s*** characterizing the dispersion of the results and given by the formula:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{2})^2}{n-1}}$$

$x_i$ being the result of the measurement $i^{th}$ and $\bar{x}$ the arithmetic mean of the ***n*** results considered.

**Repeatability standard deviation**

Standard deviation of many repetitions obtained in a single laboratory by the same operator on the same instrument, i.e. under repeatable conditions.

**Internal reproducibility standard deviation (or total intralaboratory variability)**

Standard deviation of repetitions obtained in a single laboratory with the same method, using several operators or instruments and, in particular, by taking measurements on different dates, i.e. under reproducibility conditions.

**Random error**

Result of a measurement minus the mean that would result from an infinite number of measurements of the same measurand carried out under reproducibility conditions.

**Measurement error**

Result of a measurement minus a true value of the measurand.

**Systematic error**

Mean error that would result from an infinite number of measurements of the same measurand carried out under reproducibility conditions minus a true value of the measurand.

NOTE Error is a highly theoretical concept in that it calls upon values that are not accessible in practice, in particular the true values of measurands. On principle, the error is unknown.

**Mathematical expectation**

For a series of ***n*** measurements of the same measurand, if ***n*** tends towards the infinite, the mean $\bar{x}$ tends towards the expectation ***E(x)***.

$$E(x) = n \xrightarrow{lim} \infty \frac{\sum_{i-1}^{n} x_i}{n}$$

**Calibration**

Series of operations establishing under specified conditions the relation between the values of the quantity indicated by a measuring instrument or system, or the values represented by a materialized measurement or a reference material, and the corresponding values of the quantity measured by standards.

**Intralaboratory evaluation of an analysis method**

Action which consists in submitting an analysis method to an intralaboratory statistical study, based on a standardized and/or recognized protocol, demonstrating that within its scope, the analysis method meets pre-established performance criteria.

Within the framework of this document, the evaluation of a method is based on an intralaboratory study, which includes the comparison with a reference method.

**Precision**

Closeness of agreement between independent test results obtained under prescribed conditions

Note 1 Precision depends only on the distribution of random errors and does not have any relationship with the true or specified value.

Note 2 The measurement of precision is expressed on the basis of the standard deviation of the test results.

Note 3 The expression "independent test results" refers to results obtained such that they are not influenced by a previous result on the same or a similar test material. Quantitative measurements of precision are critically dependent upon the prescribed conditions. Repeatability and reproducibility conditions are particular sets of extreme conditions.

**Quantity (measurable)**

An attribute of a phenomenon, body or substance that may be distinguished qualitatively and determined quantitatively.

**Uncertainty of measurement**

A parameter associated with the result of a measurement, which characterizes the dispersion of the values that could reasonably be attributed to the measurand.

**Standard uncertainty (*u(xi)*)**

Uncertainty of the result of a measurement expressed in the form of a standard deviation.

**Accuracy**

Closeness of agreement between the mean value obtained starting from a broad series of test results and an accepted reference value.

Note The measurement of accuracy is generally expressed in terms of bias.

**Detection limit**

Lowest amount of an analyte to be examined in a test material that can be detected and regarded as different from the blank value (with a given probability), but not necessarily quantified. In fact, two risks must be taken into account:

- the risk α of considering the substance is present in test material when its quantity is null;
- the risk β of considering a substance is absent from a substance when its quantity is not null.

**Quantification limit**

Lowest amount of an analyte to be examined in a test material that can be quantitatively determined under the experimental conditions described in the method with a defined variability (given coefficient of variation).

**Linearity**

The ability of a method of analysis, within a certain range, to provide an instrumental response or results proportional to the quality of analyte to be determined in the laboratory sample.

This proportionality is expressed by an a priori defined mathematical expression.

The linearity limits are the experimental limits of concentrations between which a linear calibration model can be applied with a known confidence level (generally taken to be equal to 1%).

**Test material**

Material or substance to which a measuring can be applied with the analysis method under consideration.

**Reference material**

Material or substance one or more of whose property values are sufficiently homogeneous and well established to be used for the calibration of an apparatus, the assessment of a measurement method, or for assigning values to materials.

**Certified reference material**

Reference material, accompanied by a certificate, one or more whose property values are certified by a procedure which establishes its traceability to an accurate realization of the unit in which the property values are expressed, and for which each certified value is accompanied by an uncertainty at a stated level of confidence.

**Matrix**

All the constituents of the test material other than the analyte.

**Analysis method**

Written procedure describing all the means and procedures required to carry out the analysis of the analyte, i.e.: scope, principle and/or reactions, definitions, reagents, apparatus, procedures, expression of results, precision, test report.

WARNING The expressions "titration method" and "determination method" are sometimes used as synonyms for the expression "analysis method". These two expressions should not be used in this way.

**Quantitative analysis method**

Analysis method making it possible to measure the analyte quantity present in the laboratory test material.

**Reference analysis method (Type I or Type II methods)**

Method, which gives the accepted reference value for the quantity of the analyte to be measured.

**Non-classified alternative method of analysis**

A routine analysis method used by the laboratory and not considered to be a reference method.

NOTE An alternative method of analysis can consist in a simplified version of the reference method.

**Measurement**

Set of operations having the object of determining a value of a quantity.

Note The operations can be carried out automatically.

**Measurand**

Particular quantity subject to measurement.

**Mean**

For a series of ***n*** measurements of the same measurand, mean value, given by the

formula:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$x_i$being the result of the $i^{th}$ measurement.

**Result of a measurement**

Value assigned to a measurand, obtained by measurement

**Sensitivity**

Ratio between the variation of the information value of the analysis method and the variation of the analyte quantity.

The variation of the analyte quantity is generally obtained by preparing various standard solutions, or by adding the analyte to a matrix.

Note 1 Defining, by extension, the sensitivity of a method as its capacity to detect small quantities should be avoided.

Note 2 A method is said to be “sensitive" if a low variation of the quantity or analyte quantity incurs a significant variation in the information value.

**Measurement signal**

Quantity representing the measurand and is functionally linked to it.

**Specificity**

Property of an analysis method to respond exclusively to the determination of the quantity of the analyte considered, with the guarantee that the measured signal comes only from the analyte.

**Tolerance**

Deviation from the reference value, as defined by the laboratory for a given level, within which a measured value of a reference material can be accepted.

**Value of a quantity**

Magnitude of a particular quantity generally expressed as a unit of measurement multiplied by a number.

**True value of a quantity**

Value compatible with the definition of a given particular quantity.

Note 1 The value that would be obtained if the measurement was perfect

Note 2 Any true value is by nature indeterminate

**Accepted reference value**

A value that serves as an agreed-upon reference for comparison and which is derived as:

a) a theoretical or established value, based on scientific principles;

b) an assigned or certified value, based on experimental work of some national or international organization;

c) a consensus or certified value, based on collaborative experimental work under the auspices of a scientific or engineering group;

Within the particular framework of this document, the accepted reference value (or conventionally true value) of the test material is given by the arithmetic mean of the values of measurements repeated as per the reference method.

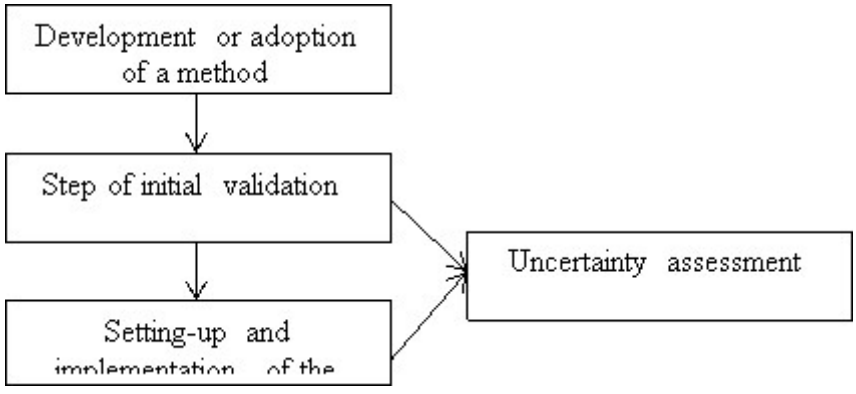
**Variance**

Square of the standard deviation.

## 4. General principles

### 4.1. Methodology

When developing a new alternative method, the laboratory implements a protocol that includes several steps. The first step, applied only once at the initial stage, or on a regular basis, is the validation of the method. This step is followed by permanent quality control. All the data collected during these two steps make it possible to assess the quality of the method. **The data collected during these two steps are used to evaluate the measurement uncertainty**. The latter, which is regularly assessed, is an indicator of the quality of the results obtained by the method under consideration.




All these steps are inter-connected and constitute a global approach that can be used to assess and control measurement errors.

### 4.2. Definition of measurement error

Any measurement carried out using the method under study gives a result which is

inevitably associated with a measurement error, defined as being the difference between the result obtained and the true value of the measurand. In practice, **the true value of the measurand is inaccessible** and a value conventionally accepted as such is used instead.

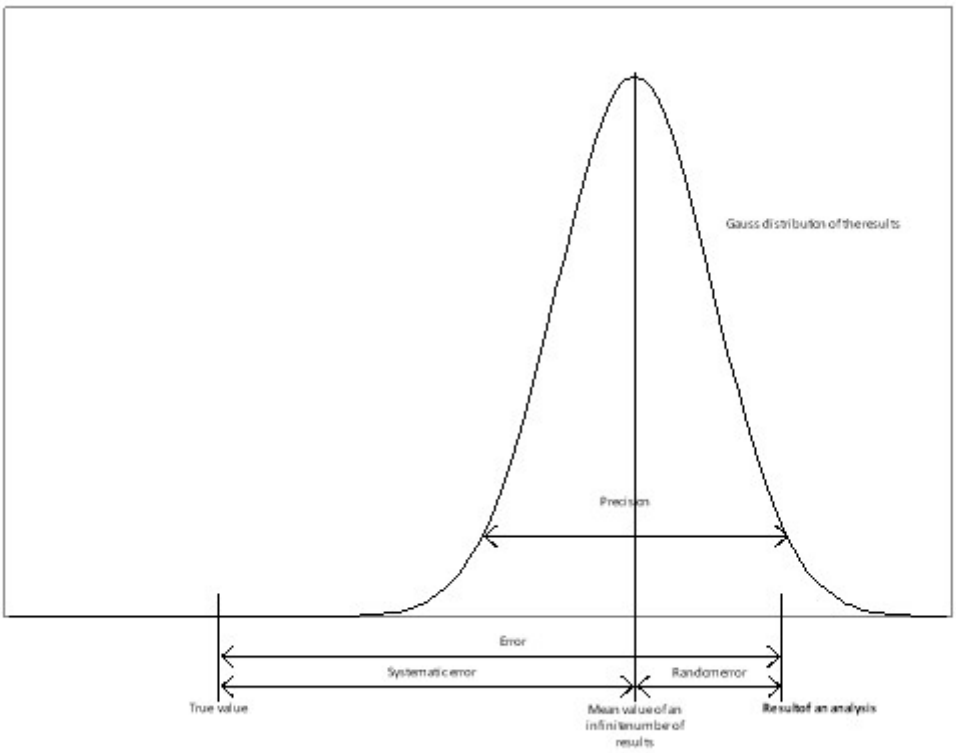
The measurement error includes two components:

Measurement error

True value= Analysis results  Systematic error  Random error

In practice, the systematic error results in a bias in relation to the true value, the random error being all the errors associated with the application of the method.

These errors can be graphically represented in the following way:




The validation and quality control tools are used to evaluate the systematic errors and the random errors, and to monitor their changes over time.

## 5. Validating a method

### 5.1. Methodology

Implementing the validation comprises 3 steps, each with objectives. To meet these objectives, the laboratory has validation tools. Sometimes there are many tools for a given objective, and are suitable for various situations. It is up to the laboratory to correctly choose the most suitable tools for the method to be validated.

| Steps | Objectives | Tools for validation |
|---|---|---|
| Scope of application | | |
| | - To define the analyzable matrices | |
| | - To define the analyzable range | Detection and quantification limit |
| | | Robustness study |
| Systematic error | | |
| or bias | - Linear response in the scale of analyzable values | Linearity study |
| | - Specificity of the method | Specificity study |
| | - Accuracy of the method | Comparison with a reference method |
| | | Comparison with reference materials |
| | | Interlaboratory comparison |

| | | |
|---|---|---|
| Random error | | |
| | - Precision of the method | Repeatability study |
| | | Intralaboratory reproducibility study |

5.2. Section one: Scope of method

5.2.1. Definition of analyzable matrices

The matrix comprises all constituents in the test material other than the analyte.

If these constituents are liable to influence the result of a measurement, the laboratory should define the matrices on which the method is applicable.

For example, in oenology, the determination of certain parameters can be influenced by the various possible matrices (wines, musts, sweet wines, etc.).

In case of doubt about a matrix effect, more in-depth studies can be carried out as part of the specificity study.

5.2.2. Detection and quantification limit

This step is of course not applicable and not necessary for those methods whose lower limit does not tend towards 0, such as alcoholic strength by volume in wines, total acidity in wines, pH, etc.

5.2.2.1. Normative definition

The detection limit is the lowest amount of analyte that can be detected but not necessarily quantified as an exact value. The detection limit is a parameter of limit tests.

The quantification limit is the lowest quantity of the compound that can be determined using the method.

5.2.2.2. Reference documents

- NF V03-110 Standard, intralaboratory validation procedure for an alternative method in relation to a reference method.
- International compendium of analysis methods – OIV, Assessment of the detection and quantification limit of an analysis method (Oeno resolution 7/2000).

5.2.2.3. Application

In practice, the quantification limit is generally more relevant than the detection limit,

the latter being by convention 1/3 of the first.

There are several approaches for assessing the detection and quantification limits:

- Determination on blank
- Approach by the linearity study
- Graphic approach

These methods are suitable for various situations, but in every case they are mathematical approaches giving results of informative value only. It seems crucial, whenever possible, to introduce a check of the value obtained, whether by one of these approaches or estimated empirically, using the checking protocol for a predetermined quantification limit.

5.2.2.4. Procedure

5.2.2.4.1. Determination on blank

5.2.2.4.1.1. Scope

This method can be applied when the blank analysis gives results with a non-zero standard deviation. The operator will judge the advisability of using reagent blanks, or matrix blanks.

If the blank, for reasons related to uncontrolled signal preprocessing, is sometimes not measurable or does not offer a recordable variation (standard deviation of 0), the operation can be carried out on a very low concentration in analyte, close to the blank.

5.2.2.4.1.2. Basic protocol and calculations

Carry out the analysis of ***n*** test materials assimilated to blanks, ***n*** being equal to or higher than 10.

- Calculate the mean of the $x_i$ results obtained:

$$\bar{x}_{blank} = \frac{\sum_{i=1}^{n} x_i}{n}$$

- Calculate the standard deviation of the $x_i$ results obtained:

$$s_{blank} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x}_{blank})^2}{n-1}}$$

- From these results the detection limit is conventionally defined by the formula:

$$L_d = \bar{x}_{blank} + (3.S_{blank})$$

- From these results the quantification limit is conventionally defined by the formula:

$$L_q = \bar{x}_{blank} + (10.S_{blank})$$

***Example:*** *The table below gives some of the results obtained when assessing the detection limit for the usual determination of free sulfur dioxide.*

| *Test material #* | *X ( mg/l)* |
|---|---|
| *1* | *0* |
| *2* | *1* |
| *3* | *0* |
| *4* | *1.5* |
| *5* | *0* |
| *6* | *1* |
| *7* | *0.5* |
| *8* | *0* |
| *9* | *0* |
| *10* | *0.5* |
| *11* | *0* |
| *12* | *0* |

*The* calculated values are as follows:

- q = 12
- $M_{blank}$ = 0.375
- $S_{blank}$ = 0.528 mg/l
- DL = 1.96 mg/l
- QL = *5.65 mg/l*

5.2.2.4.2. Approach by linearity study

5.2.2.4.2.1. Scope

This method can be applied in all cases, and is required when the analysis method does not involve background noise. It uses the data calculated during the linearity study.

Note This statistical approach may be biased and give pessimistic results when linearity is calculated on a very wide range of values for reference materials, and whose measurement results include variable standard deviations. In such cases, a linearity study limited to a range of low values, close to 0 and with a more homogeneous distribution will result in a more relevant assessment.

5.2.2.4.2.2. Basic protocol and calculations

Use the results obtained during the linearity study which made it possible to calculate the parameters of the calibration function $y = a + b.x$

The data to be recovered from the linearity study are (see chapter 5.3.1. linearity study):

- slope of the regression line:

$$b = \frac{\sum_{i-1}^{n}(x_i - M_x)(y_i - M_y)}{\sum_{i-1}^{n}(x_i - M_x)^2}$$

- residual standard deviation:

$$S_{res} = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{p}(y_{i,j} - \widehat{y_{i,j}})^2}{pm - 2}}$$

- standard deviation at the intercept point (to be calculated):

$$S_a = S_{res}\sqrt{\frac{1}{np} + \frac{M_x^2}{\sum_{i=1}^{n} p(x_i - M_x)^2}}$$

The estimates of the detection limit **DL** and the quantification limit **QL** are calculated using following formulae:

$$DL = \frac{3 \times S_a}{b}$$ Estimation detection limit

$$QL = \frac{10 \times S_a}{b}$$ Estimated quantification limit

**Example**: Estimatation of the detection and quantification limits in the determination of sorbic acid by capillary electrophoresis, based on linearity data acquired on a range from 1 to 20 $mg.L_{-1}$.

| X (ref) | Y1 | Y2 | Y3 | Y4 |
|---|---|---|---|---|
| *1* | *1.9* | *0.8* | *0.5* | *1.5* |
| *2* | *2.4* | *2* | *2.5* | *2.1* |
| *3* | *4* | *2.8* | *3.5* | *4* |
| *4* | *5.3* | *4.5* | *4.7* | *4.5* |
| *5* | *5.3* | *5.3* | *5.2* | *5.3* |
| *10* | *11.6* | *10.88* | *12.1* | *10.5* |
| *15* | *16* | *15.2* | *15.5* | *16.1* |
| *20* | *19.7* | *20.4* | *19.5* | *20.1* |

Number of reference materials
*n = 8*
Number of replicas
*p = 4*

Straight line (y = a + b*x)

*b = 0.9972*

*a = 0.51102*

residual standard deviation:
$S_{res}$ *= 0.588*

Standard deviation on the intercept point
$\boldsymbol{S}_a$ *= 0.1597*

---

The estimated detection limit is $DL = 0.48\ mg.L^{-1}$

The estimated quantification limit is $QL = 1.6\ mg.L^{-1}$

5.2.2.4.3. Graphic approach based on the background noise of the recording

5.2.2.4.3.1. Scope

This approach can be applied to analysis methods that provide a graphic recording (chromatography, etc.) with a background noise. The limits are estimated from a study of the background noise.

5.2.2.4.3.2. Basic protocol and calculation

Record a certain number of reagent blanks, using 3 series of 3 injections separated by several days.

Determine the following values:

- $h_{max}$ the greatest variation in amplitude on the y-axis of the signal observed between two acquisition points, excluding drift, at a distance equal to twenty times the width at mid-height of the peak corresponding to the analyte, centered over the retention time of the compound under study.
- R, the quantity/signal response factor, expressed in height.

The detection limit **DL**, and the quantification limit **QL** are calculated according to the following formulae:

$DL = 3\ h_{max}\ R$

$QL = 10\ h_{max}\ R$

5.2.2.4.4. Checking a predetermined quantification limit

This approach can be used to validate a quantification value obtained by statistical or empirical approach.

5.2.2.4.4.1. Scope

This method can be used to check that a given quantification limit is *a priori* acceptable. It is applicable when the laboratory can procure at least 10 test materials with known quantities of analyte, at the level of the estimated quantification limit.

In the case of methods with a specific signal, not sensitive to matrix effects, the materials can be synthetic solutions whose reference value is obtained by formulation.

In all other cases, wines (or musts) shall be used whose measurand value as obtained by the reference method is equal to the limit to be studied. Of course, in this case the quantification limit of the reference method must be lower than this value.

5.2.2.4.4.2. Basic protocol and calculation

Analyze ***n*** independent test materials whose accepted value is equal to the quantification limit to be checked; ***n*** must at least be equal to 10.

- Calculate the mean of ***n*** measurements:

$$\bar{x}_{LQ} = \frac{\sum_{i=1}^{n} x_i}{n}$$

- Calculate the standard deviation of ***n*** measurements:

$$s_{LQ} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x}_{LQ})^2}{n-1}}$$

*with $x_i$ results of the measurement of the $i^{th}$ test material.*

The two following conditions must be met:

a) the measured mean quantity $\bar{x}_{LQ}$ must not be different from the predetermined quantification limit *QL*:

If $\frac{|QL - \bar{x}QL|}{\frac{s_{QL}}{\sqrt{N}}}$ < 10 then quantification limit *QL* is considered to be valid.

Note 10 is a purely conventional value relating to the QL criterion.

b) the quantification limit must be other than 0:

If 5 $s_{QL}$ < *QL* then the quantification limit is other than 0.

A value of 5 corresponds to an approximate value for the spread of the standard deviation, taking into account risk □ and risk □ to ensure that the QL is other than 0.

This is equivalent to checking that the coefficient of variation for *QL* is lower than 20%.

Note Remember that the detection limit is obtained by dividing the quantification limit

by 3.

Note 2 A check should be made to ensure that the value of $S_{LQ}$ is not too large (which would produce an artificially positive test), and effectively corresponds to a reasonable standard deviation of the variability of the results for the level under consideration. It is up to the laboratory to make this critical evaluation of the value of $S_{LQ}$.

**Example:** Checking the quantification limit of the determination of malic acid by the enzymatic method.

Estimated quantification limit: 0.1 $g.L^{-1}$

| Wine | Values |
|---|---|
| *1* | *0.1* |
| *2* | *0.1* |
| *3* | *0.09* |
| *4* | *0.1* |
| *5* | *0.09* |
| *6* | *0.08* |
| *7* | *0.08* |
| *8* | *0.09* |
| *9* | *0.09* |
| *10* | *0.08* |

**Mean:** ***0.090***

**Standard deviation:** ***0.008***

First condition: $\frac{|LQ - \overline{x_{QL}}|}{\frac{s_{QL}}{\sqrt{n}}} = 3.87 < 10$ The quantification limit of 0.1 is considered to be valid.

Second condition: $5S_{LQ} = 0.04 < 0.1$ The quantification limit is considered to be significantly different from 0.

5.2.3. Robustness

5.2.3.1. Definition

Robustness is the capacity of a method to give close results in the presence of slight changes in the experimental conditions likely to occur during the use of the procedure.

5.2.3.2. Determination

If there is any doubt about the influence of the variation of operational parameters, the laboratory can use the scientific application of experiment schedules, enabling these critical operating parameters to be tested within the variation range likely to occur under practical conditions. In practice, these tests are difficult to implement.

5.3. Section two: systematic error study

5.3.1. Linearity study

5.3.1.1. Normative definition

The linearity of a method is its ability (within a given range) to provide an informative value or results proportional to the amount of analyte to be determined in the test material.

5.3.1.2. Reference documents

- NF V03-110 standard. Intralaboratory validation procedure of an alternative method in relation to a reference method.
- ISO 11095 Standard, linear calibration using reference materials.
- ISO 8466-1 Standard, Water quality – Calibration and evaluation of analytical methods and estimation of performance characteristics

5.3.1.3. Application

The linearity study can be used to define and validate a linear dynamic range.

This study is possible when the laboratory has stable reference materials whose accepted values have been acquired with certainty (in theory these values should have an uncertainty equal to 0). These could therefore be internal reference materials titrated with calibrated material, wines or musts whose value is given by the mean of at least 3 repetitions of the reference method, external reference materials or certified external reference materials.

In the last case, and only in this case, this study also enables the traceability of the method. The experiment schedule used here could then be considered as a

calibration.

In all cases, it is advisable to ensure that the matrix of the reference material is compatible with the method.

Lastly, calculations must be made with the final result of the measurement and not with the value of the signal.

Two approaches are proposed here:

- An ISO 11095 type of approach, the principle of which consists in comparing the residual error with the experimental error using a Fischer's test. This approach is valid above all for relatively narrow ranges (in which the measurand does not vary by more than a factor 10). In addition, under experimental conditions generating a low reproducibility error, the test becomes excessively severe. On the other hand, in the case of poor experimental conditions, the test will easily be positive and will also lose its relevance. This approach requires good homogeneity of the number of measurements over the entire range studied.
- An ISO 8466 type of approach, the principle of which consists in comparing the residual error caused by the linear regression with the residual error produced by a polynomial regression (of order 2 for example) applied to the same data. If the polynomial model gives a significantly lower residual error, a conclusion of nonlinearity could be drawn. This approach is appropriate in particular when there is a risk of high experimental dispersion at one end of the range. It is therefore naturally well-suited to analysis methods for traces. There is no need to work with a homogeneous number of measurements over the whole range, and it is even recommended to increase the number of measurements at the borders of the range.

#### 5.3.1.4. ISO 11095-type approach

##### 5.3.1.4.1. Basic protocol

It is advisable to use a number ***n*** of reference materials. The number must be higher than 3, but there is no need, however, to exceed 10. The reference materials should be measured ***p*** times, under **reproducibility conditions**, ***p*** shall be higher than 3, a number of 5 being generally recommended. The accepted values for the reference materials are to be regularly distributed over the studied range of values. The number of measurements must be identical for all the reference materials.

Note It is essential that the reproducibility conditions use a maximum of potential sources of variability, with the risk that the test shows non-linearity in an excessive way.

The results are reported in a table presented as follows:

| Reference materials | Accepted reference value material | Measured values | | | | |
|---|---|---|---|---|---|---|
| | | Replica 1 | ... | Replica j | ... | Replica p |
| *1* | $x_1$ | $y_{11}$ | ... | *y1j* | ... | *y1p* |
| ... | ... | ... | ... | ... | ... | ... |
| *i* | $x_i$ | $y_{i1}$ | ... | $y_{ij}$ | ... | *yip* |
| ... | ... | ... | ... | ... | ... | ... |
| *n* | *xn* | *yn1* | ... | *ynj* | ... | *ynp* |

5.3.1.4.2. Calculations and results

5.3.1.4.2.1. Defining the regression model

The model to be calculated and tested is as follows:

$$y_{ij} = a + b.x_i + \varepsilon_{ij}$$

where

- $y_{ij}$ is the $j^{th}$ replica of the $i^{th}$ reference material.
- $x_i$ is the accepted value of the $i^{th}$ reference material.
- b is the slope of the regression line.
- a is the intercept point of the regression line.

$a + b.x_i$ represents the expectation of the measurement value of the $i^{th}$ reference material.

$\varepsilon_{ij}$ is the difference between $y_{ij}$ and the expectation of the measurement value of the $i^{th}$ reference material.

5.3.1.4.2.2. Estimating parameters

The parameters of the regression line are obtained using the following formulae:

- mean of ***p*** measurements of the $i^{th}$ reference material

$$yi = \frac{1}{p}\sum_{j=1}^{p} y_{ij}$$

- mean of all the accepted values of *n* reference materials

$$M_x = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- mean of all the measurements

$$My = \frac{1}{n}\sum_{i=1}^{n} y_i$$

- estimated slope ***b***

$$b = \frac{\sum_{i-1}^{n}(x_i - M_x)(y_i - M_y)}{\sum_{i-1}^{n}(x_i - M_x)^2}$$

- estimated intercept point ***a***

$$a = M_y - b \times M_x$$

- regression value associated with the $i^{th}$ reference material $\hat{y}_i$

$$\hat{y}_i = a + b \times x_i$$

- residual $e_{ii}$

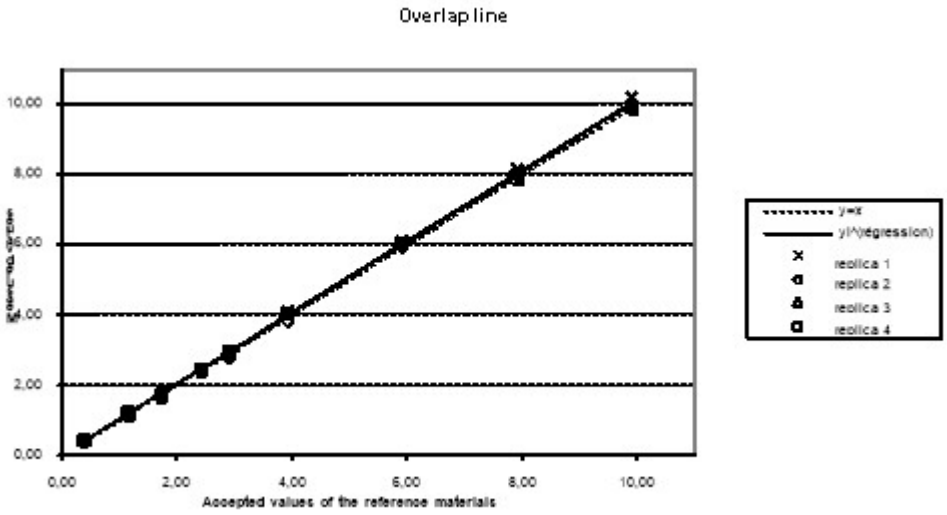
$$e_{ij} = y_{ij} - \hat{y}_i$$

5.3.1.4.2.3. Charts

The results can be presented and analyzed in graphic form. Two types of charts are used.

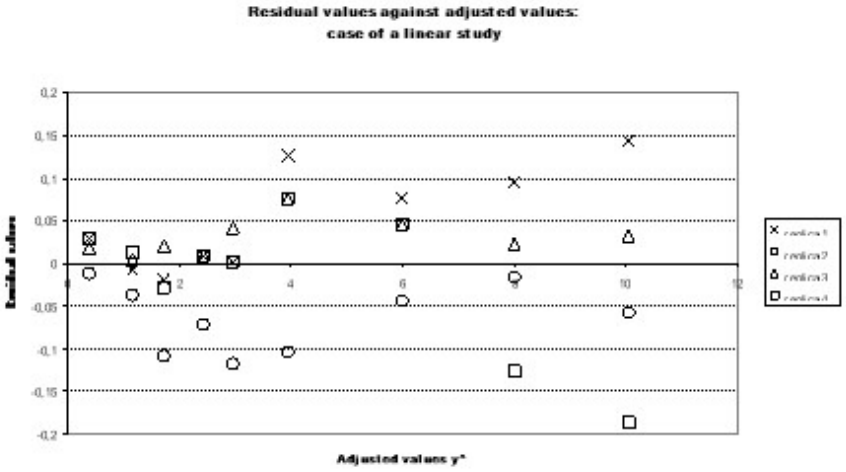
- The first type of graph is the representation of the values measured against the accepted values of reference materials. The calculated overlap line is also

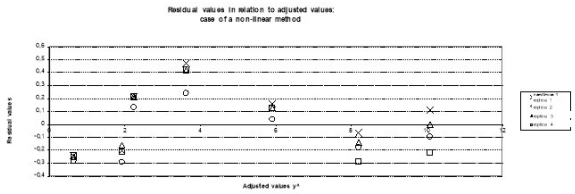
plotted.




- The second graph is the representation of the residual values against the estimated values of the reference materials ( $\hat{y}$ ) indicated by the overlap line.

The graph is a good indicator of the deviation in relation to the linearity assumption: the linear dynamic range is valid if the residual values are fairly distributed between the positive and negative values.






In case of doubt about the linearity of the regression, a Fischer-Snedecor test can be carried out in order to test the assumption: "the linear dynamic range is not valid", in addition to the graphic analysis.

5.3.1.4.2.4. Test of the linearity assumption

Several error values linked to calibration should be defined first of all: these can be estimated using the data collected during the experiment. A statistical test is then performed on the basis of these results, making it possible to test the assumption of non-validity of the linear dynamic range: this is the Fischer-Snedecor test.

Definitions of errors linked to calibration

These errors are given as a standard deviation, resulting from the square root of the ratio between a sum of squares and a degree of freedom.

**Residual error**

The residual error corresponds to the error between the measured values and the value given by the regression line.

The sum of the squares of the residual error is as follows:

$$Q_{res} = \sum_{i=1}^{n}\sum_{j=1}^{p}(y_{ij} - \hat{y}_i)^2$$

The number of degrees of freedom is *np-2*.

The residual standard deviation is then estimated by the formula:

$$S_{res} = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{p}(y_{ij} - \hat{y}_i)^2}{np - 2}}$$

**Experimental error**

The experimental error corresponds to the reproducibility standard deviation of the experimentation.

The sum of the squares of the experimental error is as follows:

$$Q_{exp} = \sum_{i=1}^{n}\sum_{j=1}^{p}(y_{ij} - y_i)^2$$

The number of degrees of freedom is *np-n*.

The experimental standard deviation (reproducibility) is then estimated by the formula:

$$S_{exp} = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{p}(y_{ij} - y_i)^2}{np - n}}$$

Note This quantity is sometimes also noted $S_R$.

**Adjustment error**

The value of the adjustment error is the experimental error minus the residual error.

The sum of the squares of the adjustment error is:

$$Q_{def} = Q_{res} - Q_{exp}$$

Or

$$Q_{def} = \sum_{i=1}^{n}\sum_{j=1}^{p}(y_{ij} - \hat{y}_i)^2 - \sum_{i=1}^{n}\sum_{j=1}^{p}(y_{ij} - y_i)^2$$

The number of degrees of freedom is n-2

The standard deviation of the adjustment error is estimated by the formula:

$$S_{def} = \sqrt{\frac{Q_{res} - Q_{exp}}{n - 2}}$$

Or

$$S_{def} = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{p}(y_{ij} - \hat{y}_i)^2 - \sum_{i=1}^{n}\sum_{j=1}^{p}(y_{ij} - y_i)^2}{n - 2}}$$

Fischer-Snedecor test

The ratio $S_{obs} = \frac{S^2_{def}}{S^2_{exp}}$ obeys the Fischer-Snedecor law with the degrees of freedom *n-2, np-n*.

The calculated experimental value $F_{obs}$ is compared with the limit value: $F_{1-\alpha}$ *(n-2,np-n)*, extracted from the Snedecor law table. The value for □ used in practice is generally 5%.

If $F_{obs} \geq F_{1-\alpha}$ the assumption of the non-validity of the linear dynamic range is accepted *(with a risk of □ error of 5%).*

If $F_{obs} \geq F_{1-\alpha}$ the assumption of the non-validity of the linear dynamic range is rejected

**Example:** Linearity study for the determination of tartaric acid by capillary electrophoresis. 9 reference materials are used. These are synthetic solutions of tartaric acid, titrated by means of a scale traceable to standard masses.

| Ref. material | Ti (ref) | Y1 | Y2 | Y3 | Y4 |
|---|---|---|---|---|---|
| *1* | *0.38* | *0.41* | *0.37* | *0.4* | *0.41* |
| *2* | *1.15* | *1.15* | *1.12* | *1.16* | *1.17* |
| *3* | *1.72* | *1.72* | *1.63* | *1.76* | *1.71* |
| *4* | *2.41* | *2.45* | *2.37* | *2.45* | *2.45* |
| *5* | *2.91* | *2.95* | *2.83* | *2.99* | *2.95* |
| *6* | *3.91* | *4.09* | *3.86* | *4.04* | *4.04* |
| *7* | *5.91* | *6.07* | *5.95* | *6.04* | *6.04* |
| *8* | *7.91* | *8.12* | *8.01* | *8.05* | *7.9* |
| *9* | *9.91* | *10.2* | *10* | *10.09* | *9.87* |

Regression line

Line ( y = a + b*x)

*b = 1.01565*

*a = - 0.00798*

Errors related to calibration

Residual standard deviation $S_{res}$ = *0.07161*

Standard deviation of experimental reproducibility $S_{exp}$ = *0.07536*

Standard deviation of the adjustment error $S_{def}$ = *0.0548*

Interpretation, Fischer-Snedecor test

$F_{obs}$ = *0.53* < $F_{1-\alpha}$ = *2.37*

***The assumption of the non-validity of the linear dynamic range is rejected***

5.3.1.5. ISO 8466-type approach

5.3.1.5.1. Basic protocol

It is advisable to use a number ***n*** of reference materials. The number must be higher than 3, but there is no need, however, to exceed 10. The reference materials should be measured several times, under **reproducibility conditions**. The number of measurements may be small at the center of the range studied (minimum = 2) and must be greater at both ends of the range, for which a minimum number of 4 is generally recommended. The accepted values of reference materials must be regularly distributed over the studied range of values.

Note It is vital that the reproducibility conditions use the maximum number of potential sources of variability.

The results are reported in a table presented as follows:

| Reference materials | Accepted value of the reference material | Measured values | | | | |
|---|---|---|---|---|---|---|
| | | Replica 1 | Replica 2 | Replica j | ... | Replica p |
| *1* | $x_1$ | $y_{11}$ | $y_{12}$ | *y1j* | ... | *y1p* |

| | | | | | | |
|---|---|---|---|---|---|---|
| ... | ... | ... | ... | ... | ... | |
| *i* | $x_i$ | $y_{i1}$ | $y_{i2}$ | | | |
| ... | ... | ... | ... | ... | ... | |
| *N* | *xn* | *yn1* | ... | *ynj* | ... | *ynp* |

5.3.1.5.2. Calculations and results

5.3.1.5.2.1. Defining the linear regression model

Calculate the linear regression model using the calculations detailed above.

The residual error of the standard deviation for the linear model $S_{res}$ can then be calculated using the formula indicated in § 5.3.1.4.2.4.1

5.3.1.5.2.2. Defining the polynomial regression model

The calculation of the polynomial model of order 2 is given below

The aim is to determine the parameters of the polynomial regression model of order 2 applicable to the data of the experiment schedule.

$$y = ax^2 + bx + c$$

The purpose is to determine the parameters a, b and c. This determination can generally be computerized using spreadsheets and statistics software.

The estimation formulae for these parameters are as follows:

$$a = \frac{\sum_i x_i^2 y_i \left(N\sum_i x_i^2 - [\sum_i xi]^2\right) - \sum_i x_i^3 \left(N\sum_i x_i y_i - \sum_i x_i \sum_i y_i\right) + \sum_i x_i^2 \left(\sum_i x_i y_i \sum_i x_i - \sum_i y_i \sum_i x_i^2\right)}{\sum_i x_i^4 \left(N\sum_I X_i^2 - [\sum_i xi]^2\right) - \sum_i x_i^3 (N\sum_i x_i^3 - \sum_i x_i^2 \sum_i xi) + \sum_i x_i^2 (\sum_i xi \sum_i x_i^3 - [\sum_i x_i^2]^2)}$$

$$b = \frac{\sum_i x_i^4 \left(N\sum_i x_i y_i - \sum_i xi \sum_i y_i\right) - \sum_i x_i^2 y_i \left(N\sum_i x_i^3 - \sum_i x_i^2 \sum_i x_i\right) + \sum_i x_i^2 \left(\sum_i y_i \sum_i x_i^3 - \sum_i x_i y_i \sum_i x_i^2\right)}{\sum_i x_i^4 \left(N\sum_I X_i^2 - [\sum_i xi]^2\right) - \sum_i x_i^3 (N\sum_i x_i^3 - \sum_i x_i^2 \sum_i xi) + \sum_i x_i^2 (\sum_i xi \sum_i x_i^3 - [\sum_i x_i^2]^2)}$$

$$c = \frac{\sum_i x_i^4 \left(\sum_i x_i^2 \sum_i y_i - \sum_i xi \sum_i x_i y_i\right) - \sum_i x_i^3 \left(\sum_i x_i^3 \sum_i y_i - \sum_i x_i^2 \sum_i x_{(y_i)}\right) + \sum_i x_i^2 y_i \left(\sum_i x_i \sum_i x_i^3 - [\sum_i x_i^2]\right)}{\sum_i x_i^4 \left(N\sum_I X_i^2 - [\sum_i xi]^2\right) - \sum_i x_i^3 (N\sum_i x_i^3 - \sum_i x_i^2 \sum_i xi) + \sum_i x_i^2 (\sum_i xi \sum_i x_i^3 - [\sum_i x_i^2]^2)}$$

Once the model has been established, the following values are to be calculated:

- Regression value associated with the $i^{th}$ reference material $\hat{y}'_i$

$$\hat{y}'_i = ax^2 + bx + c$$

- residual $e_{ij}$

$$e'_{ij} = y_{ij} - \hat{y}'_i$$

Residual standard deviation of the polynomial model

$$S'_{res} = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{p}(y_{ij} - \hat{y}'_i)^2}{np-2}}$$

Comparing residual standard deviations
Calculation of

$$DS^2 = (N-2)S^2_{res} - (N-3)S'^2_{res}$$

Then

$$PG = \frac{DS^2}{S'^2_{res}}$$

The value PG is compared with the limit value $F_{1-\alpha}$ given by the Fischer-Snedecor table for a confidence level 1- □ and a degree of freedom 1 and (N-3).

Note In general the □ risk used is 5%. In some cases the test may be optimistic and a risk of 10% will prove more realistic.

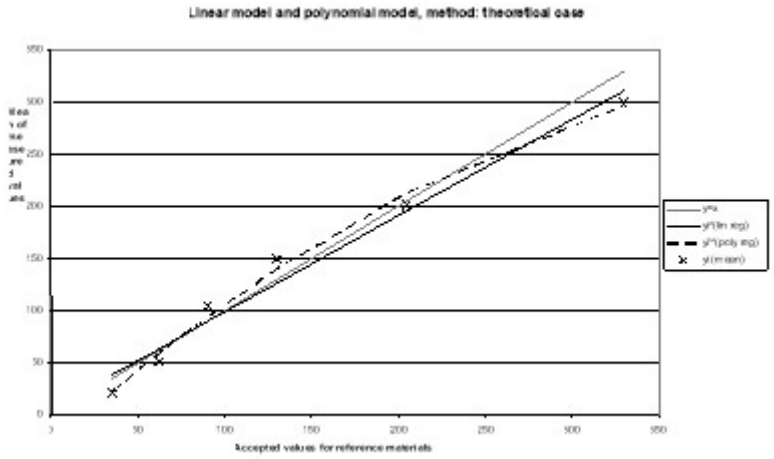
If PG $\leq$ $F_{1-\alpha}$: the nonlinear calibration function does not result in an improved adjustment; for example, the calibration function is linear.

If PG > $F_{1-\alpha}$: the work scope must be as narrow as possible to obtain a linear calibration function: otherwise, the information values from the analyzed samples must be evaluated using a nonlinear calibration function.

**Example:** Theoretical case.

| | Ti (ref) | Y1 | Y2 | Y3 | Y4 |
|---|---|---|---|---|---|
| *1* | *35* | *22.6* | *19.6* | *21.6* | *18.4* |
| *2* | *62* | *49.6* | *49.8* | *53* | |
| *3* | *90* | *105.2* | *103.5* | | |
| *4* | *130* | *149* | *149.8* | | |

| | | | | | |
|---|---|---|---|---|---|
| *5* | *205* | *203.1* | *202.5* | *197.3* | |
| *6* | *330* | *297.5* | *298.6* | *307.1* | *294.2* |




**Linear regression**

y = 1.48.x – 0.0015

$S_{res}$ = 13.625

**Polynomial regression**

y = – 0.0015x² + 1.485x – 27.2701

S'res = 7.407

**Fischer's test**

PG = 10.534 > F(5%) = 10.128

PG>F the linear calibration function cannot be retained

5.3.2. Specificity

5.3.2.1. Normative definition

The specificity of a method is its ability to measure only the compound being searched for.

5.3.2.2. Application

In case of doubt about the specificity of the tested method, the laboratory can use experiment schedules designed to check its specificity. Two types of complementary experiments are proposed here that can be used in a large number of cases encountered in the field of oenology.

- The first test is the standard addition test. It can be used to check that the

method measures all the analyte.

- The second test can be used to check the influence of other compounds on the result of the measurement.

5.3.2.3. Procedures

5.3.2.3.1. Standard addition test

5.3.2.3.1.1. Scope

This test can be used to check that the method measures all the analyte.

The experiment schedule is based on standard additions of the compound being searched for. It can only be applied to methods that are not sensitive to matrix effects.

5.3.2.3.1.2. Basic protocol

This consists in finding a significant degree of added quantities on test materials analyzed before and after the additions.

Carry out variable standard additions on ***n*** test materials. The initial concentration in analyte of test materials, and the standard additions are selected in order to cover the scope of the method. These test materials must consist of the types of matrices called for routine analysis. It is advised to use at least 10 test materials.

The results are reported in a table presented as follows:

| **Test material** | **Quantity before addition** ($x$) | **Quantity added** ($v$) | **Quantity after addition** ($w$) | **Quantity found** ($r$) |
|---|---|---|---|---|
| 1 | $x_1$ | $v_1$ | $w_1$ | $r_1 = w_1 - x_1$ |
| ... | ... | ... | ... | ... |
| $i$ | $x_i$ | $v_i$ | $w_i$ | $r_i = w_i - x_i$ |
| ... | ... | ... | ... | ... |
| $n$ | $X_n$ | $V_n$ | $w_n$ | $r_p = w_n - x_n$ |

Note 1 An addition is made with a pure standard solution. It is advised to perform an addition of the same order as the quantity of the test material on which it is carried out. This is why the most concentrated test materials must be diluted to remain within the scope of the method.

Note 2 It is advised to prepare the additions using independent standard solutions, in order to avoid any systematic error.

Note 3 The quality of values *x* and *w* can be improved by using several repetitions.

5.3.2.3.1.3. Calculations and results

The principle of the measurement of specificity consists in studying the regression line ***r = a + b.v*** and checking that slope ***b*** is equivalent to 1 and that intercept point ***a*** is equivalent to 0.

5.3.2.3.1.3.1. Study of the regression line ***r = a + b.v***

The parameters of the regression line are obtained using the following formulae:

- mean of the added quantities $\bar{v}$

$$\bar{v} = \frac{\sum_{i=1}^{n} v_i}{n}$$

- mean of the quantities found $\bar{r}$

$$\bar{r} = \frac{\sum_{i=1}^{n} r_i}{n}$$

- estimated slope ***b***

$$b = \frac{\sum_{i-1}^{n}(v_i - \bar{v})(r_i - \bar{r})}{\sum_{i-1}^{n}(v_i - \bar{v})^2}$$

- estimated intercept point ***a***

$$a = \bar{r} - b.\bar{v}$$

- regression value associated with the $i^{th}$ reference material $\hat{y}_i$

$$\hat{r}_i = a + b \times v_i$$

- residual standard deviation

$$S_{res} = \sqrt{\frac{\sum_{i=1}^{n}(r_i - \hat{r}_i)^2}{n-2}}$$

- standard deviation on the slope

$$S_b = S_{res} \sqrt{\left(\frac{1}{\sum_{i-1}^{n}(v_i - \overline{v})^2}\right)}$$

- standard deviation on the intercept point

$$S_a = S_{res} \sqrt{\frac{1}{n} + \frac{v^{-2}}{\sum_{i=1}^{n}(v_i - \overline{v})^2}}$$

5.3.2.3.1.3.2. Analysis of the results

The purpose is to conclude on the absence of any interference and on an acceptable specificity. This is true if the overlap line $r = a + bv$ is equivalent to the line $y = x$.

To do so, two tests are carried out:

- Test of the assumption that slope b of the overlap line is equal to 1.
- Test of the assumption that intercept point a is equal to 0.

These assumptions are tested using a Student test, generally associated with a risk of error of 1%. A risk of 5% can prove more realistic in some cases.

Let $T_{critical,\ bilateral}$ [dof; 1%] be a Student bilateral variable associated with a risk of error of 1% for a number of degrees of freedom (dof).

Step 1: calculations

Calculation of the comparison criterion on the slope at 1

$$T_{obs} = \frac{|b-1|}{S_b}$$

Calculation of the comparison criterion on the intercept point at 0

$$T_{obs} = \frac{|a|}{S_a}$$

Calculation of the Student critical value: $T_{critical,\ bilateral}$[ p-2; 1%]

Step 2: interpretation

- If $T_{obs}$ is lower than $T_{critical}$, then the slope of the regression line is equivalent to 1

- If $T_{obs}$ is lower than $T_{critical}$, then the intercept point of the regression line is equivalent to 0.

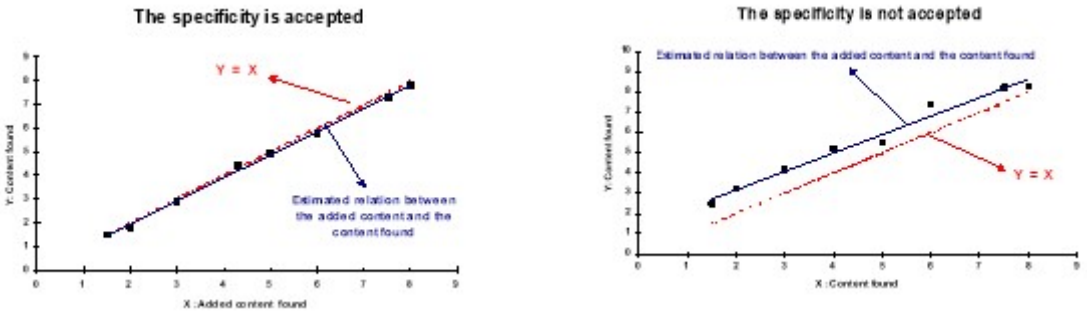
If both conditions are true, then the overlap line is equivalent = *y* = *x*, and the method is deemed to be specific.

Note 1 Based on these results, a mean overlap rate can be calculated to quantify the specificity. In no case should it be used to "correct" the results. This is because if a significant bias is detected, the alternative method cannot be validated in relation to an efficiency rate of 100%.

Note 2 Since the principle of the test consists in calculating a straight line, at least three levels of addition have to be taken, and their value must be correctly chosen in order to obtain an optimum distribution of the points.

5.3.2.3.1.3.3. Overlap line graphics

**Example of specificity**




5.3.2.3.2. Study of the influence of other compounds on the measurement result

5.3.2.3.2.1. Scope

If the laboratory suspects the interaction of compounds other than the analyte, an experiment schedule can be set up to test the influence of various compounds. The experiment schedule proposed here enables a search for the influence of compounds defined *a priori*: thanks to its knowledge of the analytical process and its know-how, the laboratory should be able to define a certain number of compounds liable to be present in the wine and to influence the analytical result.

5.3.2.3.2.2. Basic protocol and calculations

Analyze ***n*** wines in duplicate, before and after the addition of the compound suspected of having an influence on the analytical result; ***n*** must at least be equal to 10.

The mean values ***Mxi*** of the 2 measurements $x_i$ and $x'_i$ made before the addition shall be calculated first, then the mean values ***My***$_i$ of the 2 measurements $y_i$ and $y'_i$ made after the addition, and finally the difference $d_i$ between the values $Mx_i$ and $My_i$.
The results of the experiment can be reported as indicated in the following table:

| | ***x***: Before addition | | ***y***: After addition | | Means | | Difference |
|---|---|---|---|---|---|---|---|
| **Samples** | **Rep1** | **Rep2** | **Rep1** | **Rep2** | ***x*** | ***y*** | ***d*** |
| *1* | $x_1$ | $x'_1$ | $y_1$ | $y'_1$ | $Mx_1$ | $My_1$ | $d_1 = Mx_1 - My_1$ |
| ... | ... | ... | ... | ... | ... | ... | ... |
| *i* | $x_i$ | $x'_i$ | $y_i$ | $y'_i$ | $Mx_i$ | $My_i$ | $d_i = Mxi - My_i$ |
| ... | ... | ... | ... | ... | ... | ... | ... |
| *n* | $x_n$ | $x'_n$ | $y_n$ | $y'_n$ | $Mx_n$ | $My_n$ | $d_n = Mx_n - My_n$ |

The mean of the results before addition $M_x$

$$M_x = \frac{1}{n}\sum_{i=1}^{n} Mx_i$$

The mean of the results after addition $M_y$

$$M_y = \frac{1}{n}\sum_{i=1}^{n} My_i$$

Calculate the mean of the differences $M_d$

$$M_d = \sum_{i=1}^{n} \frac{d_i}{n} = M_y - M_x$$

Calculate the standard deviation of the differences $S_d$

$$S_d = \sqrt{\frac{\sum_{i=1}^{n}(d_i - M_d)^2}{n-1}}$$

Calculate the Z-score

$$Z_{score} = \frac{|M_d|}{S_d}$$

5.3.2.3.2.3. Interpretation

If the $Z_{score}$ is $\leq$ 2, the added compound can be considered to have a negligible influence on the result of analysis with a risk of 5%.

If the $Z_{score}$ is $\geq$ 2, the added compound can be considered to influence the result of analysis with a risk of 5%.

Note Interpreting the $Z_{score}$ is possible given the assumption that the variations obey a normal law with a 95% confidence rate.

**Example:** Study of the interaction of compounds liable to be present in the samples, on the determination of fructose glucose in wines by Fourier transform infrared spectrophotometry (FTIR).

| | Before addition | | + 250 mg.L$^{-1}$ potassium sorbate | | + 1 g. L$^{-1}$ salicylic acid | | Differences | |
|---|---|---|---|---|---|---|---|---|
| vin | rep1 | rep2 | rep1 | rep2 | rep1 | rep2 | sorbate diff | salicylic diff |
| *1* | *6.2* | *6.2* | *6.5* | *6.3* | *5.3* | *5.5* | *0.2* | *-0.8* |
| *2* | *1.2* | *1.2* | *1.3* | *1.2* | *0.5* | *0.6* | *0.05* | *-0.65* |
| *3* | *0.5* | *0.6* | *0.5* | *0.5* | *0.2* | *0.3* | *-0.05* | *-0.3* |
| *4* | *4.3* | *4.2* | *4.1* | *4.3* | *3.8* | *3.9* | *-0.05* | *-0.4* |
| *5* | *12.5* | *12.6* | *12.5* | *12.7* | *11.5* | *11.4* | *0.05* | *-1.1* |
| *6* | *5.3* | *5.3* | *5.4* | *5.3* | *4.2* | *4.3* | *0.05* | *-1.05* |
| *7* | *2.5* | *2.5* | *2.6* | *2.5* | *1.5* | *1.4* | *0.05* | *-1.05* |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *8* | *1.2* | *1.3* | *1.2* | *1.1* | *0.5* | *0.4* | *-0.1* | *-0.8* |
| *9* | *0.8* | *0.8* | *0.9* | *0.8* | *0.2* | *0.3* | *0.05* | *-0.55* |
| *10* | *0.6* | *0.6* | *0.5* | *0.6* | *0.1* | *0* | *-0.05* | *-0.55* |

| | | | |
|---|---|---|---|
| Potassium sorbate | *Md =* | *0.02* | |
| | *Sd =* | *0.086* | |
| | $Z_{score}$ = | *0.23* | *<2* |
| Salicylic acid | *Md =* | *-0.725* | |
| | *Sd =* | *0.282* | |
| | $Z_{score}$ = | *2.57* | *>2* |

In conclusion, it can be stated that potassium sorbate does not influence the determination of fructose glucose by the FTIR gauging studied here. On the other hand, salicylic acid has an influence, and care should be taken to avoid samples containing salicylic acid, in order to remain within the scope of validity for the gauging under study.

5.3.3. Study of method accuracy

5.3.3.1. Presentation of the step

5.3.3.1.1. Definition

Correlation between the mean value obtained with a large series of test results and an accepted reference value.

5.3.3.1.2. General principles

When the reference value is output by a certified system, the accuracy study can be regarded a traceability link. This applies to two specific cases in particular:

- Traceability to certified reference materials: in this case, the accuracy study can be undertaken jointly with the linearity and calibration study, using the experiment schedule described for that study.
- Traceability to a certified interlaboratory comparison analysis chain.

The other cases, i.e. which use references that are not based on certified systems, are the most widespread in routine oenological laboratories. These involve comparisons:

- Comparison with a reference method
- Comparison with the results of an uncertified interlaboratory comparison analysis chain.
- Comparison with internal reference materials, or with external uncertified reference materials.

5.3.3.1.3. Reference documents

- NF V03-110 Standard. intralaboratory validation procedure for an alternative method in relation to a reference method.
- NF V03-115 Standard, Guide for the use of reference materials.
- ISO 11095 Standard, linear calibration using reference materials.
- ISO 8466-1 Standard. Water quality – Calibration and evaluation of analytical methods and estimation of performance characteristics
- ISO 57025 Standard, Exactitude of results and methods of measurement

5.3.3.2. Comparison of the alternative method with the OIV reference method

5.3.3.2.1. Scope

This method can be applied if the laboratory uses the OIV reference method, or a traced, validated method, whose performance quality is known and meets the requirements of the laboratory's customers.

To study the comparative accuracy of the two methods, it is advisable first of all to ensure the quality of the repeatability of the method to be validated, and to compare it with the reference method. The method for carrying out the repeatability comparison is described in the chapter on repeatability.

5.3.3.2.2. Accuracy of the alternative method compared with the reference method

5.3.3.2.2.1. Definition

Accuracy is defined as the closeness of agreement between the values obtained by the reference method and that obtained by the alternative method, independent of the errors of precision of the two methods.

5.3.3.2.2.2. Scope

The accuracy of the alternative method in relation to the reference method is established for a field of application in which the repeatabilities of the two methods are constant.

In practice, it is therefore often advisable to divide the analyzable range of values into several sections or "range levels" (2 to 5), in which we may reasonably consider that the repeatabilities of the methods are comparable to a constant.

5.3.3.2.2.3. Basic protocol and calculations

In each range level, accuracy is based on a series of ***n*** test materials with concentration values in analyte covering the range level in question. A minimum number of 10 test materials is required to obtain significant results.

Each test material is to be analyzed in duplicate by the two methods under repeatable conditions.

A calculation is to be made of the mean values $Mx_i$ of the 2 measurements $x_i$ et $x'_i$ made using the alternative method and the mean values $My_i$ of the 2 measurements $y_i$ et $y'_i$ made using the reference method, then the difference $d_i$ is to be calculated between the values $Mx_i$ and $My_i$.

The results of the experiment can be reported as in the following table:

| | ***x***: Alternative method | | ***y:*** Reference method | | Means | | Difference |
|---|---|---|---|---|---|---|---|
| **Test material** | **Rep1** | **Rep2** | **Rep1** | **Rep2** | ***x*** | ***y*** | ***d*** |
| *1* | $x_1$ | $x'_1$ | $y_1$ | $y'_1$ | $Mx_1$ | $My_1$ | $d_1 = Mx_1 - My_1$ |
| *...* | *...* | *...* | *...* | *...* | *...* | *...* | *...* |
| *i* | $x_i$ | $x'_i$ | $y_i$ | $y'_i$ | $Mx_i$ | $My_i$ | $d_i = Mxi - My_i$ |
| *...* | *...* | *...* | *...* | *...* | *...* | *...* | *...* |
| *n* | $x_n$ | $x'_n$ | $y_n$ | $y'_n$ | $Mx_n$ | $My_n$ | $d_n = Mx_n - My_n$ |

The following calculations are to be made

- The mean of the results for the alternative method $M_x$

$$M_x = \frac{1}{n}\sum_{i=1}^{n} Mx_i$$

- The mean of the results for the reference method $M_y$

$$M_y = \frac{1}{n}\sum_{i=1}^{1} My_i$$

- Calculate the mean of the differences $M_d$

$$M_d = \sum_{i=1}^{n} \frac{d_i}{n} = M_x - M_y$$

- Calculate the standard deviation of the differences $S_d$

$$S_d = \sqrt{\frac{\sum_{i=1}^{n}(d_i - M_d)^2}{n-1}}$$

- Calculate the $Z_{score}$

$$Z_{score} = \frac{|M_d|}{S_d}$$

5.3.3.2.2.4. Interpretation

- If the $Z_{score}$ is **lower** than or equal to 2.0, it can be concluded that the accuracy of one method in relation to the other is satisfactory, in the range level under consideration, with a risk of error α = 5%.
- If the $Z_{score}$ is **higher** than 2.0, it can be concluded that the alternative method is not accurate in relation to the reference method, in the range level under consideration, with a risk of error α = 5%.

Note Interpreting the $Z_{score}$ is possible given the assumption that the variations obey a normal law with a 95% confidence rate.

**Example:** Study of the accuracy of FTIR gauging to determine glucose and fructose in

relation to the enzymatic method. The first range level covers the scale from 0 to 5 $g.L^{-1}$ and the second range level covers a scale from 5 to 20 $g.L^{-1}$.

| Wine | FTIR 1 | IRTF2 | Enz 1 | Enz 2 | di |
|---|---|---|---|---|---|
| *1* | *0* | *0.3* | *0.3* | *0.2* | *-0.1* |
| *2* | *0.2* | *0.3* | *0.1* | *0.1* | *0.2* |
| *3* | *0.6* | *0.9* | *0.0* | *0.0* | *0.7* |
| *4* | *0.7* | *1* | *0.8* | *0.7* | *0.1* |
| *5* | *1.2* | *1.6* | *1.1* | *1.3* | *0.2* |
| *6* | *1.3* | *1.4* | *1.3* | *1.3* | *0.0* |
| *7* | *2.1* | *2* | *1.9* | *2.1* | *0.0* |
| *8* | *2.4* | *0* | *1.1* | *1.2* | *0.1* |
| *9* | *2.8* | *2.5* | *2.0* | *2.6* | *0.3* |
| *10* | *3.5* | *4.2* | *3.7* | *3.8* | *0.1* |
| *11* | *4.4* | *4.1* | *4.1* | *4.4* | *0.0* |
| *12* | *4.8* | *5.4* | *5.5* | *5.0* | *-0.2* |

| | | | | | |
|---|---|---|---|---|---|
| *Md* | *0.13* | | | | |
| *Sd* | *0.23* | | | | |
| $Z_{score}$ | *0.55* | *< 2* | | | |

| Wine | FTIR 1 | IRTF2 | Enz 1 | Enz 2 | di |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| *1* | *5.1* | *5.4* | *5.1* | *5.1* | *0.1* |
| *2* | *5.3* | *5.7* | *5.3* | *6.0* | *-0.2* |
| *3* | *7.7* | *7.6* | *7.2* | *7.0* | *0.6* |
| *4* | *8.6* | *8.6* | *8.3* | *8.5* | *0.2* |
| *5* | *9.8* | *9.9* | *9.1* | *9.3* | *0.6* |
| *6* | *9.9* | *9.8* | *9.8* | *10.2* | *-0.1* |
| *7* | *11.5* | *11.9* | *13.3* | *13.0* | *-1.4* |

For the two range levels, the $Z_{score}$ is lower than 2. The FTIR gauging for the determination of fructose glucose studied here, can be considered accurate in relation to the enzymatic method.

5.3.3.3. Comparison by interlaboratory tests

5.3.3.3.1. Scope

Interlaboratory tests are of two types:

- **Collaborative studies** relate to a single method. These studies are carried out for the initial validation of a new method, mainly in order to define the standard deviation of interlaboratory reproducibility $SR_{inter}$***(method)***. The mean ***m*** could also be given.
- Interlaboratory comparison analysis chains, or **aptitude tests**. These tests are carried out for the validation of a method adopted by the laboratory, and the routine quality control (see § 5.3.3.3). The resulting value is the interlaboratory mean ***m***, as well as the standard interlaboratory reproducibility and intermethod deviation ***SRinter***.

By participating in an analysis chain, or in a collaborative study, the laboratory can exploit the results in order to study the accuracy of a method, in order to ensure its validation first of all, and its routine quality control.

If the interlaboratory tests are carried out within the framework of a certified organization, this comparison work can be used for method traceability.

5.3.3.3.2. Basic protocol and calculations

To obtain a sufficient comparison, it is recommended to use a minimum number of 5 test materials over the period.

For each test material, two results are provided:

- The mean of all the laboratories with significant results ***m***
- The standard deviation for interlaboratory reproducibility $S_{R-inter}$

The test materials are analyzed with ***p*** replicas by the laboratory, these replicas being carried out under repeatable conditions. ***p*** must at least be equal to 2.

In addition, the laboratory must be able to check that the intralaboratory variability (intralaboratory reproducibility) is lower than the interlaboratory variability (interlaboratory reproducibility) given by the analysis chain.

For each test material, the laboratory calculates the $\boldsymbol{Z_{score}}$, given by the following formula:

$$Z_{score} = \frac{|m_{lab} - m|}{S_{R-inter}}$$

The results can be reported as indicated in the following table:

| Test material | Rep1 | ... | Rep j | ... | Rep p | Lab mean | Chain mean | Standard deviation | $Z_{score}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | $x_{11}$ | ... | $x_{1j}$ | ... | $x_{1p}$ | $m_{labi} = \frac{\sum_{j=1}^{p} x_{ij}}{p}$ | $m_1$ | $S_{R\text{-inter}(1)}$ | $Z_{score} = \frac{\lvert m_{labi} - m_i \rvert}{S_{R-inter(i)}}$ |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| *i* | $x_{i1}$ | ... | $x_{ij}$ | ... | $x_{ip}$ | $m_{labi} = \frac{\sum_{j=1}^{p} x_{j}}{p}$ | $m_i$ | $S_{R\text{-inter}(i)}$ | $Z_{score} = \frac{\lvert m_{labi} - m_i \rvert}{S_{R-inter(i)}}$ |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| n | $x_{n1}$ | ... | $x_{nj}$ | ... | $x_{np}$ | $m_{labi} = \frac{\sum_{j=1}^{p} x_{nj}}{p}$ | $m_n$ | $S_{R\text{-inte}(n)}$ | $Z_{score} = \frac{\lvert m_{labi} - m_n \rvert}{S_{R-inter(n)}}$ |

5.3.3.3.3. Interpretation

If all the $Z_{score}$ results are lower than 2, the results of the method being studied can considered identical to those obtained by the laboratories having produced significant results.

Note Interpreting the $Z_{score}$ is possible given the assumption that the variations obey a normal law with a 95% confidence rate.

**Example:** An interlaboratory analysis chain outputs the following results for the free

sulfur dioxide parameter, on two samples.

| Samples | $x_1$ | $x_2$ | $x_3$ | $x_4$ | Lab mean | Chain mean | Standard deviation | $Z_{score}$ |
|---|---|---|---|---|---|---|---|---|
| *1* | *34* | *34* | *33* | *34* | *33.75* | *32* | *6* | *0.29* <2 |
| *2* | *26* | *27* | *26* | *26* | *26.25* | *24* | *4* | *0.56* <2 |

It can be concluded that on these two samples, the comparison with the analysis chain is satisfactory.

5.3.3.4. Comparison with reference materials

5.3.3.4.1. Scope

In situations where there is no reference method (or any other method) for a given parameter, and the parameter is not processed by the analysis chains, the only remaining possibility is comparison of the results of the method to be validated with accepted internal or external material reference values.

The reference materials, for example, could be synthetic solutions established with class-A glassware, and/or calibrated metrology apparatus.

In the case of certified reference materials, the comparison constitutes the traceability value, and can be carried out at the same time as the gauging and linearity study.

5.3.3.4.2. Basic protocol and calculations

It is advisable to have ***n*** reference materials for a given range level, in which it can be reasonably estimated that repeatability is comparable to a constant; ***n*** must at least be equal to 10.

Analyze in duplicate each reference material.

Calculate the mean values $Mx_i$ for the 2 measurements $x_i$ and $x'_i$ carried out using the alternative method.

Define $T_i$ the accepted value for the $i^{th}$ reference material.

The results can be reported as indicated in the following table:

| **Reference material** | ***x***: Alternative method | | | ***T***: Accepted value of the reference material | Difference |
|---|---|---|---|---|---|
| | **Rep1** | **Rep2** | **Mean x** | | ***d*** |
| *1* | $x_1$ | $x'_1$ | $Mx_1$ | $T_1$ | $d_{1=}Mx_1-T_1$ |

| | | | | | |
|---|---|---|---|---|---|
| ... | | | ... | ... | ... |
| *i* | $x_i$ | $x'_i$ | $Mx_i$ | $T_i$ | $d_i = Mxi-T_i$ |
| ... | | | ... | ... | ... |
| *n* | $x_n$ | $x'_n$ | $Mx_n$ | $T_n$ | $d_n = Mx_n-T_n$ |

The mean of the results of the alternative method $M_x$

$$M_x = \frac{1}{n}\sum_{i=1}^{n} Mx_i$$

The mean of the accepted values of reference materials $M_T$

$$M_T = \frac{1}{n}\sum_{i=1}^{n} T_i$$

Calculate the mean of the differences $M_d$

$$M_d = \sum_{i=1}^{n} \frac{d_i}{n} = Mx - M_T$$

Calculate the standard deviation of the differences $S_d$

$$S_d = \sqrt{\frac{\sum_{i=1}^{n}(d_i - M_d)^2}{n-1}}$$

Calculate the Z-score

$$Z_{score} = \frac{|M_d|}{S_d}$$

5.3.3.4.3. Interpretation

- If the $Z_{score}$ is **lower than** or equal to 2.0, it can be concluded that the accuracy of the alternative method in relation to the accepted values for the reference material is good on the range level under consideration.
- If $Z_{score}$ is **higher** than 2.0, it can be concluded that the alternative method is not accurate in relation to the accepted values for the reference materials in the

range level under consideration.

Note Interpreting the $Z_{score}$ is possible given the assumption that the variations obey a normal law with a 95% confidence rate.

**Example:** There is no reference method to compare the results of the analysis of Ethyl-4 Phenol (4-EP) by Gas chromatography coupled with mass spectrometry (GC-MS). The results are compared with the accepted values for reference materials, consisting of synthetic solutions formulated by traced equipment.

| Test apparatus | Ti (ref) | Y1 | Y2 | Y3 | Y4 | My | $d_i$ |
|---|---|---|---|---|---|---|---|
| *1* | *4.62* | *6.2* | *6.56* | *4.9* | *5.7* | *5.8* | *1.2* |
| *2* | *12.3* | *15.1* | *10.94* | *12.3* | *11.6* | *12.5* | *0.2* |
| *3* | *24.6* | *24.5* | *18* | *25.7* | *27.8* | *24.0* | *-0.6* |
| *4* | *46.2* | *48.2* | *52.95* | *46.8* | *35* | *45.7* | *-0.5* |
| *5* | *77* | *80.72* | *81.36* | *83.2* | *74.5* | *79.9* | *2.9* |
| *6* | *92.4* | *97.6* | *89* | *94.5* | *99.5* | *95.2* | *2.8* |
| *7* | *123.2* | *126.6* | *129.9* | *119.6* | *126.9* | *125.8* | *2.6* |
| *8* | *246.4* | *254.1* | *250.9* | *243.9* | *240.4* | *247.3* | *0.9* |
| *9* | *385* | *375.8* | *366.9* | *380.4* | *386.9* | *377.5* | *-7.5* |
| *10* | *462* | *467.5* | *454.5* | *433.3* | *457.3* | *453.2* | *-8.9* |

*Md = -0.7*

*Sd = .4.16*

$Z_{score}$ *= 0.16*

Given these results, the values obtained by the analysis method for 4-EP by GC-MS can be considered accurate compared with the accepted values of reference materials.

5.4. Section three: random error study

5.4.1. General principle

Random error is approximated using precision studies. Precision is calculated used a methodology that can be applied under various experimental conditions, ranging between those of repeatability, and those of reproducibility, which constitute the extreme conditions of its measurement.

The precision study is one of the essential items in the study of the uncertainty of measurement.

5.4.2. Reference documents

- ISO 5725 Standard, Exactitude of results and methods of measurement
- NF V03-110 Standard, Intralaboratory validation procedure for an alternative method in relation to a reference method.

5.4.3. Precision of the method

5.4.3.1. Definition

Closeness of agreement between independent test results obtained under prescribed conditions.

Note 1 Precision depends only on the distribution of the random errors and has no relation with the true or specified value.

Note 2 Expressing the measurement of precision is based on the standard deviation of the test results.

Note 3 The term "independent test results" refers to results obtained such that they are not influenced by a previous result on the same or similar test material. Quantitative measurements of precision are critically dependent on the prescribed conditions. Repeatability and reproducibility conditions are particular sets of extreme conditions.

In practice, precision refers to all the experimental conditions ranging between the conditions of repeatability and those of reproducibility.

5.4.3.2. Scope

The protocols and calculations are detailed below, from the general theoretical case to the specific cases of repeatability and reproducibility. This exhaustive approach should make it possible to apply the precision study in most laboratory situations.

The precision study can be applied a priori without difficulty to every quantitative method.

In many cases, precision is not constant throughout the validity range for the method. In this case, it is advisable to define several sections or "range levels", in which we may reasonably consider that the precision is comparable to a constant. The calculation of

precision is to be reiterated for each range level.

5.4.3.3. General theoretical case

5.4.3.3.1. Basic protocol and calculations

5.4.3.3.1.1. Calculations with several test materials

***n*** test materials are analyzed over a relatively long period of time with several replicas, $p_i$ being the number of replicas for the $i^{th}$ test material. The properties of the test materials must maintain constant throughout the period in question.

For each replica, the measurement can be made with ***K*** repetitions, (we do not take into account the case here where the number of repetitions ***K*** can vary from one test material to the other, which would complicate the calculations even more).

The total number of replicas must be higher than 10, distributed over all the test materials.

The results can be reported as indicated in the following table, (case in which ***K*** = 2)

| **Replicas** | **1** | | **...** | | ***j*** | | $p_1$ | | $p_i$ | | $p_n$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Test materials.* | | | | | | | | | | | | |
| 1 | $x_{11}$ | $x'_{11}$ | ... | ... | $x_{1j}$ | $x'_{1j}$ | $x_{1p1}$ | $x'_{1p1}$ | | | | |
| ... | | | | | | | | | | | | |
| *i* | $x_{i1}$ | $x'_{i1}$ | ... | ... | $x_{ij}$ | $x'_{ij}$ | ... | ... | $x_{ipi}$ | $x'_{ipi}$ | | |
| ... | | | | | | | | | | | | |
| *n* | $x_{n1}$ | $x'_{n1}$ | ... | ... | $x_{nj}$ | $x'_{nj}$ | ... | ... | ... | ... | $x_{npn}$ | $x'_{npn}$ |

In this situation, the standard deviation of total variability (or standard deviation of precision $S_v$) is given by the general expression:

$$s_v = \sqrt{Var(\bar{x}_{ij}) + \left(1 - \frac{1}{k}\right) Var(répet)}$$

where:

$Var(\bar{x}_{ij})$ variance of the mean of repeated replicas of all test materials.

$Var(répet)$ variance of the repeatability of all the repetitions.

- If the test materials were analyzed in duplicate with each replica (***K*** **=** 2), the expression becomes:

$$S_v = \sqrt{Var(\bar{x}_{ij}) + \frac{Var(repeat)}{2}}$$

- When only one measurement of the test material has been carried out with each replica (***K*** **=** 1), the variance of repeatability is null, the expression becomes:

$$s_v = \sqrt{Var(\bar{x}_{ij}}$$

- **Calculation of** $Var(\bar{x}_{ij})$

The mean of the two replicas $x_{ij}$ and $x'_{ij}$ is:

$$\bar{x}_{ij} = \frac{x_{ij} + x'_{ij}}{2}$$

For each test material, the mean of ***n*** replicas is calculated:

$$Mx_i = \frac{\sum_{j-1}^{p_i} \bar{x}_{ij}}{p_i}$$

The number of different measurements ***n*** is the sum of $p_i$

$$N = \sum_{i=1}^{n} p_i$$

The variance $Var(\bar{x}_{ij})$ is then given by the following equation

$$Var(x_{ij}) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{p_i} (\bar{x}_{ij} - Mx_j)^2}{N - n}$$

Note This variance can also be calculated using the variances of variability of each test material: $Var_i$ ***(x$_i$)***. The following relation is then used (it is strictly equivalent to the previous one):

$$Var(x_{ij}) = \frac{\sum_{i=1}^{n} (p_i - 1).Var_i(x_j)}{N - n}$$

- **Calculation of** $Var(repeat)$

The variance of repeatability is calculated as a conventional repeatability equation with ***n*** test materials in duplicate. According to the calculation of repeatability discussed in the section entitled "repeatability", for ***K = 2*** the variance of repeatability is:

$$Var(repeat) = \frac{\sum_{i-1}^{p} \sum_{j-1}^{n_i} w_{ij}^2}{2N} \qquad where\ w_{ij} = x_{ij} - x'_{ij}$$

Precision ***v*** is calculated according to the formula:

$$v = 2\sqrt{2}.S_v = 2.8.S_v$$

The value of precision ***v*** means that in 95% of the cases, the difference between two values obtained by the method, under the conditions defined, will be lower than or equal to ***v***.

Note 1 The use and interpretation of these results is based on the assumption that the variations obey a normal law with a 95% confidence rate.

Note 2 One can also define a precision of 99% with $v=2.58\sqrt{2}.S_v=3.65.S_v$

5.4.3.3.1.2. Calculations with 1 test material

In this situation, the calculations are simpler. It is advisable to carry out ***p*** measurement replicas of the test material, if necessary with a repetition of the measurement on each replica. ***p*** must at least be equal to 10.

In the following calculations, the measurement is considered to be carried out in duplicate with each replica.

- The variance $Var(\bar{x}_{ij})$ is then given by the following equation:

$$Var(x_{ij}) = \frac{\sum_{i-1}^{p} (\bar{x}_i - M_x)^2}{p - 1}$$

where:

$\bar{x}_i$ is the mean of the two repetitions of replica ***i***

***p*** is the number of replicas

$M_x$ is the mean of all the replicas

- The variance $Var(repeat)$ is then given by the following equation:

$$Var(repeat) = \frac{\sum_{i-1}^{p} w_i^2}{2p}$$

*where* $w_i$*: difference between the two repetitions of replica* ***i***

5.4.3.4. Repeatability

5.4.3.4.1. Definitions

Repeatability is the closeness of agreement between mutually-independent analysis results obtained with the method in question on the same wine, in the same laboratory, with the same operator using the same equipment, within a short period of time.

These experimental conditions will be called conditions of repeatability.

The value of repeatability **r** is the value below which the absolute difference between two results of the same analysis is considered to be located, obtained under the conditions of repeatability defined above, with a confidence level of 95%.

The repeatability standard deviation **Sr** is the standard deviation for the results obtained under the conditions of repeatability. It is a parameter of the dispersion of the results, obtained under conditions of repeatability.

5.4.3.4.2. Scope

A priori, the repeatability study can be applied without difficulty to every quantitative method, insofar as the repeatability conditions can be observed.

In many cases, repeatability is not constant throughout the range of validity of the method. It is therefore advisable to define several sections or "range levels", in which we may reasonably consider that the repeatability is comparable to a constant. The repeatability calculation is then to be reiterated for each range level.

5.4.3.4.3. Basic protocol and calculations

5.4.3.4.3.1. General case

The number of test materials may vary in relation to the NUMBER of replicas. In practice, we consider that the number of measurements of all test materials must be higher than 20. It is not necessary for the repeatability conditions to be maintained from one test material to another, but all the replicas carried out on the same test material must be carried out under these repeatability conditions.

Repeatability remains a special case of the precision calculation $s_v = \sqrt{Var(\bar{x}_{ij}) + \frac{Var(repeat)}{2}}$. The $Var(repeat)$ part is naturally equal to 0 (only one measurement with each replica), and the calculation is the same as the calculation of $Var\ (x_{ij})$

$$s_r = \sqrt{Var(x_{ij})} = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{p_i}(\bar{x}_{ij} - M_{x_i})^2}{N-n}}$$

The value ***r*** means that in 95% of the cases, the difference between two values acquired under repeatable conditions will be lower than or equal to ***r***.

5.4.3.4.3.2. Particular case applicable to only 1 repetition

In practice, the most current situation for automated systems is the analysis of test material with only one repetition. It is advisable to use at least 10 materials in order to reach the 20 measurements required. The two measurement replicas of the same test material must be carried out under repeatable conditions.

In this precise case, the calculation of $S_r$ is simplified and becomes:

$$S_r = \sqrt{\frac{\sum_{i=1}^{q} w_i^2}{2p}}$$

in which:

$s_r$ = the repeatability standard deviation

***p*** = the number of test materials analyzed in duplicate

$w_i$ = the absolute differences between duplicates

Repeatability **r** is calculated according to the formula:

***r = 2.8 Sr***

**Example:** For the alternative determination method of the free sulfur dioxide in question, and for a range of measurements from 0 to 50 mg/l, the operator will seek at least 10 samples with regularly distributed concentrations ranging between these values.

| *Sample no.* | $x_i$ *(in mg/l)* | $x'_i$ *(in mg/l)* | $W_i$ *(absolute value)* |
|---|---|---|---|
| *1* | *14* | *14* | *0* |
| *2* | *25* | *24* | *1* |
| *3* | *10* | *10* | *0* |

| | | | |
|---|---|---|---|
| *4* | *2* | *3* | *1* |
| *5* | *35* | *35* | *0* |
| *6* | *19* | *19* | *0* |
| *7* | *23* | *23* | *0* |
| *8* | *27* | *27* | *0* |
| *9* | *44* | *45* | *1* |
| *10* | *30* | *30* | *0* |
| *11* | *8* | *8* | *0* |
| *12* | *48* | *46* | *2* |

**Example:** Using the values given in the table above, the following results are obtained:

**Q** = 12

**Sr** = 0.54 mg/l

**R** = 1.5 mg/l

This result can be used to state that, with a probability of 95%, the results obtained by the method under study will have a repeatability rate lower than 1.5 mg/l.

5.4.3.4.4. Comparison of repeatability

5.4.3.4.4.1. Determination of the repeatability of each method

To estimate the performance of a method, it can be useful to compare its repeatability with that of a reference method.

Let $S_{r-alt}$ be the repeatability standard deviation of the alternative method, and $S_{r-ref}$ the repeatability standard deviation of the reference method.

The comparison is direct. If the value of repeatability of the alternative method is lower than or equal to that of the reference method, the result is positive. If it is higher, the laboratory must ensure that the result rests compliant with the specification that it accepted for the method concerned. In the latter case, it may also apply a Fischer-Snedecor test to know if the value found for the alternative method is significantly higher than that of the reference method.

5.4.3.4.4.2. Fischer-Snedecor test

Calculate the ratio:

$$F_{obs} = \frac{S^2_{r_{alt}}}{S^2_{r_{ref}}}$$

Use the critical Snedecor value with a risk □ equal to 0.05 corresponding to the Fischer variable with a confidence level 1 □, in which □1 = n(x)-n, and □2 = n(z)-m degrees of freedom: F(N(x)-n, N(y)-m, 1- □). In the case of a calculated repeatability with only one repetition on *p* test materials for the alternative method, and *q* test materials for the reference method, the Fischer variable will have as a degree of freedom □1 = *p*, and □2 = *Q*, i.e.: F(p, Q, 1- □).

Interpreting the test:

1/ $F_{obs} > F_{1-\alpha}$ the repeatability value of the alternative method is significantly higher than that of the reference method.

2/ $F_{obs} < F_{1-\alpha}$, we cannot state that the repeatability value of the alternative method is significantly higher than that of the reference method.

***Example:*** *The value of the repeatability standard deviation found for the determination method of free sulfur dioxide is:*

***Sr*** *= 0.54 mg/l*

*The laboratory carried out the determination on the same test materials using the OIV reference method. The value of the repeatability standard deviation found in this case is:*

***Sref*** *= 0.39 mg/l*

$$F_{obs} = \frac{0.54^2}{0.39^2} = \frac{0.29}{0.15} = 1.93$$

$\boldsymbol{V_2}$ = *12*

$\boldsymbol{V_1}$ = *12*

$\boldsymbol{F_{1-\alpha}}$ = *2.69 > 1.93*

*The* $F_{obs}$ *value obtained is lower than the value* $\boldsymbol{F_{1-\alpha}}$*; we cannot state that the repeatability value of the alternative method is significantly higher than that of the reference method.*

5.4.3.5. Intralaboratory reproducibility

5.4.3.5.1. Definition

Intralaboratory reproducibility is the closeness of agreement between the analysis results obtained with the method under consideration on the same wine, in the same laboratory, with the same operator or different operators using from the different gauging curves, on different days.

5.4.3.5.2. Scope

Reproducibility studies can be implemented on quantitative methods, if the time of analysis is reasonably limited, and if the capacity exists to keep at least one test material stable over time.

In many cases, reproducibility is not constant throughout the validity range of the method. In this case, it is advisable to define several sections or "range levels", in which it can be reasonably considered that reproducibility is comparable to a constant. The reproducibility calculation is then to be reiterated for each range level.

5.4.3.5.3. Basic protocol and calculations

The laboratory chooses one or more stable test materials. It applies the method regularly for a period equal to at least one month and keeps the results obtained ($X_{ij}$ **material *i*, replica *j***). A minimum of 5 replicas is recommended for each test material, the total minimum number of replicas being 10. The replicas can be analyzed in duplicate.

The calculation of precision fully applies to the calculation of reproducibility, integrating $Var(repeat)$ if the measurements are carried out in duplicate.

Reproducibility ***R*** is calculated according to the formula:

$R = 2.8\ S_R$

The value ***R*** means that in 95% of the cases, the difference between two values acquired under reproducibility conditions will be lower than or equal to ***R***.

**Example:** Reproducibility study of the determination of the sorbic acid in wines by steam distillation and reading by absorption at 256 Nm.

Two different sorbated wines were kept for a period of 3 months. The determination of the sorbic acid was carried out at regular intervals over this period, with repetition of each measurement.

| | Test material 1 | | Test material 2 | |
|---|---|---|---|---|
| Replicas | **x1** | **x2** | **x1** | **x2** |

| | | | | |
|---|---|---|---|---|
| *1* | *122* | *125* | *140* | *139* |
| *2* | *123* | *120* | *138* | *137* |
| *3* | *132* | *130* | *139* | *141* |
| *4* | *121* | *115* | *143* | *142* |
| *5* | *130* | *135* | *139* | *139* |
| *6* | *135* | *142* | *135* | *138* |
| *7* | *137* | *135* | *139* | *139* |
| *8* | *130* | *125* | *145* | *145* |
| *9* | *123* | *130* | *138* | *137* |
| *10* | *112* | *115* | *135* | *134* |
| *11* | *131* | *128* | *146* | *146* |
| *12* | | | *137* | *138* |
| *13* | | | *146* | *147* |
| *14* | | | *145* | *148* |
| *15* | | | *130* | *128* |

$n = 2$

$p_1 = 11$

$p_2 = 15$

$n = 26$

$Var(x_{ij}) = 37.8$

$Var(repet) = 5.01$

$\boldsymbol{S_R = 6.35}$

$\boldsymbol{R = 17.8}$

## 6. Quality control of analysis methods (IQC)

### 6.1. Reference documents

- Resolution OIV Œno 19/2002: Harmonized recommendations for internal quality control in analysis laboratories.
- CITAC/EURACHEM: Guide for quality in analytical chemistry, 2002 Edition
- Standard NF V03-115, Guide for the use of reference materials

### 6.2. General principles

It is recalled that an analysis result can be affected two types of error: systematic error, which translates into bias, and random error. For series analyses, another type of error can be defined, which can be due to both systematic error and random error: this is the series effect, illustrated for example by the deviation of the measuring system during a series.

The IQC is designed to monitor and control these three errors.

### 6.3. Reference materials

The IQC is primarily based on exploiting the measurement results for reference materials. The choice and constitution of the materials are therefore essential steps that it must be controlled in order to provide an efficient basis for the system.

*A reference material is defined by two parameters:*

- *Its matri*
- *The assignment of its reference value*

Several cases are possible; the cases encountered in oenology are summarized in the following two-dimensional table:

| | | | | |
|---|---|---|---|---|
| **Doped wine**<br>*A doped wine is a wine with an artificial addition of an analyte.* | This method is applicable when the base wine is completely free of analyte. These types of materials are suitable for oenological additives that are not native to the wine. If doping is applied with a component native to the wine, the matrix can no longer be considered natural. Doping must be carried out according to metrological rules. The value obtained is prone to uncertainty.<br>*This case can be used to monitor the precision of the method, as well as its accuracy in a point. It can be applied to methods sensitive to matrix effects for non-native compounds of the wine, but not in the case of native compounds of the wine.* | In practice, this involves conditioned wine samples doped and/or chemically stabilized as proposed by organizations. These materials cannot claim to constitute a natural matrix. The reference values are generally generated by an analysis chain.<br>*This case can be used to monitor the precision of the method, as well as its accuracy in a point compared with the external standard. This has a traceability value in this point if the organization supplying the samples has been approved for the preparation of the reference material in question. It cannot be applied to methods sensitive to matrix effects.* | The measurement is carried out 3 times with the reference method, the value retained is the mean of the 3 results, insofar as they remain within an interval lower than the repeatability of the method.<br>*This case can be used to monitor the precision of a method, and to check its accuracy in a point compared with the reference method. It can be applied to methods sensitive to matrix effects for non-native compounds of the wine, but not in the case of native compounds of the wine.* | The reference value is measured using the method to be checked. The material is measured over 10 repetitions, and a check is made to ensure that the differences between these values are lower than the repeatability value; the most extreme values can be withdrawn, up to a limit of two values withdrawn. To ensure the consistency of the values obtained during the 10 repetitions, the series should be checked using control materials established during a previous session, placed at the start and end of the series.<br>*This case can only be used to monitor the precision of the method; accuracy must be monitored using another method. It can be applied to methods sensitive to matrix effects for non-native compounds of the wine, but not in the case of native compounds of the wine.* |

| | | | | |
|---|---|---|---|---|
| **Natural matrix (wine etc.)** *Natural matrices a priori constitute the most interesting reference materials because they avoid any risk of matrix effect for methods that are not perfectly specific.* | Not applicable | The external value has been determined on the wine by an interlaboratory analysis chain. Certain organizations propose conditioned wine samples whose values have been determined in this way. However, in certain cases, the wines presented in this way may have been doped and/or chemically stabilized, which means the matrix may be affected. *This case can be used to monitor the precision of a method, and to check its accuracy in a point compared with the external value. This has traceability value in this point if the analysis chain has been accredited. It can be applied to methods sensitive to matrix effects.* | The measurement is carried out 3 times with the reference method, the selected value is the mean of the 3 results, insofar as they remain within an interval lower than the repeatability of the method. *This case can be used to monitor the precision of a method, and to check its accuracy in a point compared with the reference method. It can be applied to methods sensitive to matrix effects.* | The reference value is measured by the method to be checked. The material is measured over 10 repetitions, and a check is to be made that the differences between these values are lower than the repeatability value; the most extreme values can be withdrawn, up to a limit of two values. To ensure the consistency of the values obtained over the 10 repetitions, this series is to be checked on the one hand by control materials established during a previous session, placed at the start and end of the series. The value obtained can also be compared with the value obtained by the reference method (during the 3 repetitions for example). The difference between the two values must remain lower than the calculated accuracy of the alternative method compared with the reference method. *This case is of interest in particular when a method produces a random reproducible error specific to each sample, in particular because of the non-specificity of the measured signal. This error is often minimal and lower than the uncertainty, but can generate a systematic error if the method is adjusted on a single value. This can be used to monitor the precision of the method, accuracy must be monitored using another approach. This is notably the case of the FTIR.* |

| | | | | |
|---|---|---|---|---|
| **Synthetic solution**<br>*Synthetic solutions can be used to constitute reference materials quite easily. They are not compatible with methods with non-specific signals, and that are sensitive to matrix effects.* | The solution must be produced using metrological rules. It is recalled that the formulation value obtained is prone to uncertainty.<br>*The application of such a case can be used to monitor the precision of the method, as well as its accuracy in a point in relation to a calibrated reference.* | The organization supplying the solution must provide guarantees about its quality and be certified if possible. The reference values will be accompanied by an uncertainty value at a given confidence level.<br>*This case can be used to monitor the precision of a method, and to check its accuracy in a point compared with the external value. This has traceability value in this point if the supplier organization is approved for the preparation of reference material in question. It cannot be applied to methods sensitive to matrix effects.* | If the synthetic solution has not been obtained with a calibrated material, the reference value can be determined by analyzing the synthetic solution using the reference method. The measurement is to be carried out at least 3 times. The selected value is the mean of the 3 results, insofar as they remain within an interval lower than the repeatability of the method. If necessary, the operator can check the consistency of the results obtained with the formulation value for the solution.<br>*This case can be used to monitor the precision of a method, and to check its accuracy in a point compared with the reference method. It cannot be applied to methods sensitive to matrix effects.* | The reference value is measured by the method to be checked. The material is measured over 10 repetitions, and a check will be made that the differences between these values are lower than the repeatability value; the most extreme values can be withdrawn, up to a limit of two values. To ensure the consistency of the values obtained over the 10 repetitions, the series is to be checked using control materials established during a previous session, placed at the start and end of the series.<br>*This case can be used to monitor only the precision of the method, accuracy must be monitored using another approach.* |
| **Matrix**<br><br>**Reference value** | ***Value obtained by formulation*** | ***External value to the laboratory*** | ***Value obtained by a reference method*** | ***Value obtained by the method to be checked***<br>*The use of the instrument value as a reference value does not control accuracy. An alternative approach must be set up.* |

6.4. Checking the analytical series

6.4.1. Definition

An analytical series is a series of measurements carried out under repeatable conditions.

For a laboratory that mainly uses the analytical series method of analysis, a check must be made to ensure the instantaneous adjustment of the measuring instrument and its stability during the analytical series is correct.

Two complementary approaches are possible:

- the use of reference materials (often called by extension "control materials")
- the use of an internal standard, in particular for separative methods.

6.4.2. Checking accuracy using reference materials

*Systematic error can be checked by introducing reference materials, the reference value of which has been assigned using means external to the method being checked.*

*The measured value of the reference material is associated with a tolerance limit, inside which the measured value is accepted as being valid. The laboratory defines tolerance values for each parameter and for each analytical system.* ***These values are specific to the laboratory****.*

*The control materials must be selected so that their reference values correspond to the levels of the values usually found for a given parameter. If the scale of measurement is broad, and the uncertainty of measurement is not constant on the scale, several control materials should be used to cover the various range levels.*

6.4.3. Intraseries precision

*When the analytical series are rather long, there is a risk of drift of the analytical system. In this case, intraseries precision must be checked using the same reference material positioned at regular intervals in the series. The same control materials as those used for accuracy can be used.*

*The variation in the measured values for same reference material during the series should be lower than the repeatability value* r *calculated for a confidence level of 95%.*

*Note For a confidence level of 99%, a value of 3.65.*$S_r$ *can be used.*

6.4.4. Internal standard

Certain separative methods enable the introduction of an internal standard into the product to be analyzed.

In this case, an internal standard should be introduced with calibrated material with a known uncertainty of measurement.

The internal standard enables a check to be made both of intraseries accuracy and precision. It should be noted that a drift affects the signals of the analyte and of the internal standard in equal proportions; since the value of the analyte is calculated with

the value of the signal of the internal standard, the effect of the drift is cancelled.

The series will be validated if the internal standards are inside the defined tolerance values.

6.5. Checking the analysis system

6.5.1. Definition

This concerns an additional check to the series check. It differs from the latter in that it compiles values acquired over long time scales, and/or compares them with values resulting from other analysis systems.

Two applications will be developed:

- Shewhart charts to monitor the stability of the analysis system
- Internal and external comparison of the analysis system

6.5.2. Shewhart chart

Shewhart charts are graphic statistical tools used to monitor the drift of measurement systems, by the regular analysis, in practice under reproducibility conditions, of stable reference materials.

6.5.2.1. Data acquisition

A stable reference material is measured for a sufficiently long period, at defined regular intervals. These measurements are recorded and logged in control charts. The measurements are made under reproducibility conditions, and are in fact exploitable for the calculation of reproducibility, and for the assessment of measurement uncertainty.

The values of the analytical parameters of the reference materials selected must be within valid measurement ranges.

The reference materials are analyzed during an analytical series, routine if possible, with a variable position in the series from one time to another. In practice, it is perfectly possible to use the measurements of control materials of the series to input the control charts.

6.5.2.2. Presentation of results and definition of limits

The individual results are compared with the accepted value of the reference material, and with the reproducibility standard deviation for the parameter in question, at the range level in question.

Two types of limits are defined in the Shewhart charts, the limits associated with individual results, and the limits associated with the mean.

The limits defined for the individual results are usually based on the standard

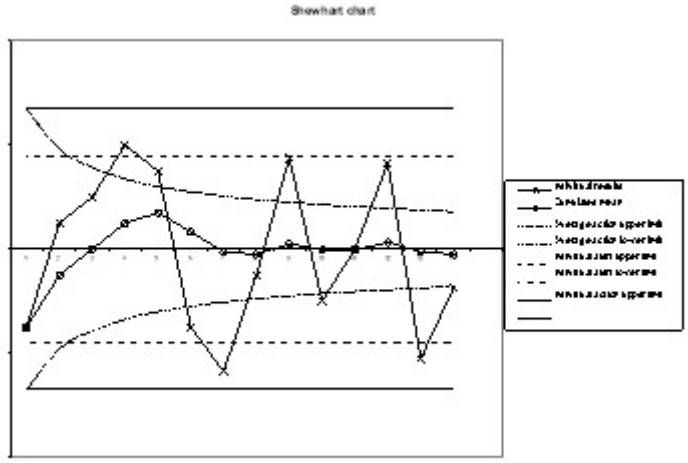
deviation values for intralaboratory reproducibility for the range level in question. They are of two types:

- alert limit: $+/-2.S_R$
- action limit: $+/-3.S_R$.

The limit defined for the cumulated mean narrows as the number of measurements increases.

This limit is an action limit: $+/-\frac{3.S_R}{\sqrt{n}}$. ***n*** being the number of measurements indicated on the chart.

Note For reasons of legibility, the alert limit of the cumulated mean is only rarely reproduced on the control chart, and has as its value $+/-\frac{2.S_R}{\sqrt{n}}$.




6.5.2.3. Using the Shewhart chart

Below we indicate the operating criteria most frequently used. It is up to the laboratories to precisely define the criteria they apply.

Corrective action on the method (or the apparatus) will be undertaken:

a) if an individual result is outside the action limits of the individual results.

b) if two consecutive individual results are located outside the alert limits of individual results.

c) if, in addition, *a posteriori* analysis of the control charts indicates a drift in the method in three cases:

- nine consecutive individual result points are located on the same side of the line

of the reference values.

- six successive individual result points ascend or descend.
- two successive points out of three are located between the alert limit and the action limit.

d) if the arithmetic mean of ***n*** recorded results is beyond one of the action limits of the cumulated mean (which highlights a systematic deviation of the results).

*Note The control chart must be revised at* **n** = *1 as soon as a corrective action has been carried out on the method.*

6.5.3. Internal comparison of analysis systems

In a laboratory that has several analysis methods for a given parameter, it is interesting to carry out measurements of the same test materials in order to compare the results. The agreement of the results between the two methods is considered to be satisfactory if their variation remains lower than 2 times the standard deviation of difference calculated during validation, with a confidence level of 95%.

Note This interpretation is possible given the assumption that the variations obey a normal law with a 95% confidence rate.

6.5.4. External comparison of the analysis system

6.5.4.1. Analysis chain of interlaboratory comparisons

The organization of the tests and calculations is given in the chapter "comparison with an interlaboratory analysis chain".

In addition to checking the accuracy by the $Z_{score}$ the results can be analyzed in greater detail, in particular with regard to the position of the values of the laboratory in relation to the mean. If they are systematically on the same side of the mean for several successive analysis chains, this can justify the implementation of corrective action by the laboratory, even if $Z_{score}$ remains lower than the critical value.

Note Interpreting the $Z_{score}$ is possible given the assumption that the variations obey a normal law with a 95% confidence rate.

If the intercomparison chain is subject to accreditation, this work of comparison has traceability value.

6.5.4.2. Comparison with external reference materials

Measuring external reference materials at regular intervals also can be used to supervise the occurrence of a systematic error (bias).

The principle is to measure the external reference material, and to accept or refuse

the value in relation to tolerance limits. These limits are defined in relation to the combination of the uncertainties of the controlled method and the reference value of the reference material.

6.5.4.2.1. Standard uncertainty of reference material

The reference values of these materials are accompanied by confidence intervals. The laboratory must determine the nature of this data, and deduce from them the standard uncertainty value for the reference value $S_{ref}$ A distinction must be made between several cases:

- The case in which uncertainty ***a*** is given in the form of an interval confidence at 95% (expanded uncertainty). This means that a normal law has been adopted. ***a*** therefore constitutes an "expanded uncertainty" and corresponds to 2 times the standard deviation $S_{ref}$ of the standard uncertainty of the reference values of the materials provided.

$$S_{ref} = \frac{a}{2}$$

The case of a certificate, or another specification, giving limits +/- ***a*** without specifying the confidence level. In this case, a rectangular dispersion has been adopted, and the value of measurement X has the same chance of having an unspecified value in the interval ***ref+/- a***.

$$S_{ref} = \frac{a}{\sqrt{3}}$$

The particular case of glassware giving limits +/- ***a***. This is the framework of a triangular dispersion.

$$S_{ref} = \frac{a}{\sqrt{6}}$$

6.5.4.2.2. Defining the validity limits of measuring reference material

To standard uncertainty $S_{ref}$ of the value of the external reference material, is added the standard uncertainty of the laboratory method to be checked,$S_{method}$. These two sources of variability must be taken into account in order to determine the limits.

$S_{method}$is calculated from the expanded uncertainty of the laboratory method in the following way:

$$S_{method} = \frac{uncerta\ int\ y}{2}$$

The validity limit of the result (with a confidence level of 95%) =

$$reference\ value +/- 2\sqrt{S_{ref}^2 + S_{method}^2}$$

**Example:** A pH 7 buffer solution is used to check a pH-meter. The confidence interval given by the pH solution is +/- 0.01. It is indicated that this confidence interval corresponds to the expanded uncertainty with a confidence level of 95%. In addition the expanded uncertainty of the pH-meter is 0.024.

The limits will be

$$+/- 2\sqrt{(\frac{0.01}{2})^2 + (\frac{0.024}{2})^2}$$

i.e. **+/- 0.026** in relation to the reference value, with a confidence level of 95%.

## 7. Assessment of measurement uncertainty

### 7.1. Definition

Parameter, associated with the result of a measurement, which characterizes the dispersion of the values that can reasonably be allotted to the measurand.

In practice, uncertainty is expressed in the form of a standard deviation called standard uncertainty ***u(x)***, or in an expanded form (generally with k = 2) ***U = +/- k.u***

### 7.2. Reference documents

- AFNOR ENV 13005 Standard: 1999 – Guide for expressing measurement uncertainty
- EURACHEM, 2000. Quantifying Uncertainty in Analytical Measurement, *EURACHEM second edition 2000*
- ISO 5725 Standard: 1994 – Exactitude (accuracy and precision) of results and measurement methods
- ISO 21748 standard: 2004 – Guidelines relating to the use of estimations of repeatability, reproducibility and accuracy in evaluating measurement uncertainty
- Perruchet C and Priel M., Estimating uncertainty, *AFNOR Publications, 2000*

### 7.3. Scope

Uncertainty provides two types of information.

- On the one hand, that intended for the customers of the laboratory, indicating the potential variations to take into account in order to interpret the result of an analysis. It must be indicated, however, that this information cannot be used as an external means of evaluating the laboratory.
- In addition, it constitutes a dynamic in-house tool for evaluating the quality of the laboratory analysis results. Insofar as its evaluation is regular and based on a fixed, well-defined methodology, it can be used to see whether the variations involved in a method change positively or negatively (in the case of an estimate based exclusively on intralaboratory data).

The present guide limits itself to providing a practical methodology for oenological laboratories dealing with series analyses. These laboratories have large volumes of data of a significant statistical scale.

Estimating uncertainties can therefore be carried out in most cases using the data collected as part validation and quality control work (in particular with the data in the Shewhart charts). These data can be supplemented by experiment schedules, in particular to determine the systematic errors.

The reference systems describe two main approaches for determining uncertainty: the intralaboratory approach and the approach interlaboratory. Each provides results that are naturally and significantly different. Their significance and their interpretation cannot be identical.

- **the intralaboratory approach** provides a result specific to the method in question, in the laboratory in question. The uncertainty that results is an indicator of the performance of the laboratory for the method in question. It answers the customer as follows: "what dispersion of results can I expect from the laboratory practicing the method?"
- **the interlaboratory approach** uses results resulting from interlaboratory tests, which provide information about the overall performance of the method.

Laboratories can use the two approaches jointly. It will be interesting to see whether the results obtained using the intralaboratory approach give values lower than the values of the interlaboratory approach.

7.4. Methodology

The work of uncertainty assessment involves 3 fundamental steps.

- Definition of the measurand, and description of the quantitative analysis method
- Critical analysis of the measurement process

- Uncertainty assessment.

7.4.1. Definition of the measurand, and description of the quantitative analysis method

First of all, the following must be specified:

- the purpose of the measurement
- the quantity measured
- If the measurand is to be obtained by calculation based on measured quantities, if possible the mathematical relation between them should be stipulated.
- all the operating conditions.

These items are included in theory in the procedures of the laboratory quality system.

In certain cases the expression of the mathematical relation between the measurand and the quantities can be highly complex (physical methods etc.), and it is neither necessarily relevant nor possible to fully detail them.

7.4.2. Critical analysis of the measurement process

The sources of error influencing the final result should be identified in order to constitute the uncertainty budget. The importance of each source can be estimated, in order to eliminate those that have only a negligible minor influence. This is done by estimating:

- the degree of gravity of the drift generated by poor control of the factor in question
- the frequency of the potential problems
- their detectability.

This critical analysis can, for example, be carried out using the "5M" method.

Labor;

Operator effect

Matter:

Sample effect (stability, homogeneity, matrix effects), and consumables (reagents, products, solutions, reference materials), etc.

Hardware:

Equipment effect (response, sensitivity, integration modes, etc.), and laboratory equipment (balance, glassware etc.).

Method:

Application effect of the procedure (operating conditions, succession of the operations etc.).

Medium:

Environmental conditions (temperature, pressure, lighting, vibration, radiation, moisture etc.).

7.4.3. Estimation calculations of standard uncertainty (intralaboratory approach)

7.4.3.1. Principle

In the case of laboratories using large series of samples with a limited number of methods, a statistical approach based on intralaboratory reproducibility, supplemented by the calculation of sources of errors not taken into account under intralaboratory reproducibility conditions, appears to be the most suitable approach.

An analysis result deviated from the true value under the effect of two sources of error: systematic errors and random errors.

**Analysis result = True value + Systematic error + Random error**

Uncertainty characterizes the dispersion of the analysis result. This translates into a standard deviation.

Variability (analysis result) = uncertainty

Variability (true value) = 0

Variability (systematic error) = $\sqrt{\sum S^2_{systematic_errors}}$

Variability (random error) = $S_R$ (intralaboratory reproducibility standard deviation)

Since standard deviations are squared when added, the estimated standard uncertainty ***u(x)*** takes the following form:

$$u(x) = \sqrt{\sum u^2_{(systematic_errors} + S^2_R}$$

Non-integrable sources of errors under the intralaboratory reproducibility conditions, i.e. systematic errors, must be determined in the form of standard deviation to be combined together and with the reproducibility standard deviation.

The laboratory can take action so that the reproducibility conditions applied make it possible to include a maximum number of sources of errors. This is obtained in particular by constituting stable test materials over a sufficiently long period, during which the laboratory takes care to vary all the possible experimental factors. In this

way, $S_R$ will cover the greatest number of possible sources of errors (random), and the work involved in estimating the systematic errors, which is often more complex to realize, will be minimized.

It should be noted here that the EURACHEM/CITAC guide entitled "Quantifying uncertainty in analytical measurements" recalls that "In general, the ISO Guide requires that corrections be applied for all systematic effects that are identified and significant". In a method "under control", systematic errors should therefore constitute a minor part of uncertainty.

The following non-exhaustive table gives examples of typical sources of error and proposes an estimation approach for each of them, using integration under reproducibility conditions as much as possible.

| *Source of error* | *Type of error* | *Commentary* | *Estimation method* |
|---|---|---|---|
| Sampling (constitution of the sample) | Random | Sampling is one of the "businesses" defined in the ISO 17025 standard. Laboratories stating they do not perform sampling, do not include this source of error in the uncertainty assessment. | Can be including in intralaboratory reproducibility by including sampling in handling. |
| Sub-sampling (sampling a quantity of sample in order to carry out the test) | Random | Is significant if the sample is not homogeneous. This source of error remains minor for wine. | Included in the intralaboratory reproducibility conditions if the test material used is similar to routine test materials. |

| | | | |
|---|---|---|---|
| Stability of the sample | Random | Depends on the storage conditions of the sample. In the case of wines, laboratories should pay detailed attention to the losses of sulfur dioxide and ethanol. | Possible changes in the sample can be integrated into the reproducibility conditions. This source of uncertainty can then be evaluated overall. |
| Gauging of the apparatus | Systematic/Random<br>This error is systematic if gauging is established for a long period, and becomes random if gauging is regularly carried out over a time-scale integrated under reproducibility conditions | Source of error to be taken into account in absolute methods. | Error of gauging line § 7.4.2.4.1<br>Taken into account under the reproducibility conditions if gauging is regularly revised. |
| Effect of contamination or memory | Random | This effect will be minimized by the proper design of measuring instruments and suitable rinsing operations | The reproducibility conditions take this effect into account, as long as the reference materials are inserted at various positions in the analysis series. |

| | | | |
|---|---|---|---|
| Precision of automata | Random | This applies to intraseries drift in particular. This can be controlled in particular by positioning the control materials within the framework of the IQC | The reproducibility conditions take this effect into account, as long as the reference materials are inserted at various positions in the analysis series. |
| Purity of the reagents | Random | The purity of the reagents has very little effect on the relative methods, insofar as the gauging and analyses are carried out with the same batches of reagents. This effect is to be taken into account in absolute methods. | To be integrated under reproducibility conditions using various batches of reagents. |
| Purity of the reagents | Random | The purity of the reagents has very little effect on the relative methods, insofar as the gauging and analyses are carried out with the same batches of reagents. This effect is to be taken into account in absolute methods. | To be integrated under reproducibility conditions using various batches of reagents. |
| Measurement conditions | Random | Effects of temperature, moisture etc. | Typically taken into account under reproducibility conditions |

| | | | |
|---|---|---|---|
| Matrix effect | Random from one sample to another, systematic on the same sample | These effects are to be taken into account in methods whose measured signal is not perfectly specific. | If this effect is regarded as significant, a specific experiment schedule can be used to estimate uncertainty due to this effect § 7.4.2.4.3 This effect is not integrated under reproducibility conditions. |
| Gauging effect | Systematic if gauging is constant Random if gauging is regularly renewed | | Taken into account under the reproducibility conditions if gauging is regularly renewed. If the gauging used remains the same one (on the scale of the periods in question within the framework of the reproducibility conditions), it is advisable to implement an experiment schedule in order to estimate the error of the gauging line § 7.4.2.4.1 |
| Operator effect | Random | | To be taken into account in the reproducibility conditions by taking care to utilize all the authorized operators. |

| | | | |
|---|---|---|---|
| Bias | Systematic | Must be minimized by the quality control work of the laboratory. | Systematic effect, can be estimated using certified references. |

7.4.3.2. Calculating the standard deviation of intralaboratory reproducibility

The reproducibility standard deviation $S_R$ is calculated using the protocol described in the section entitled "Intralaboratory reproducibility" (cf. § 5.4.3.5).

The calculation can be based on several test materials. In the noteworthy case where $S_R$ is proportional to the size of the measurand, the data collected on several test materials with different values should not be combined: $S_R$ should be expressed in relative value (%).

7.4.3.3. Estimating typical sources of systematic errors not taken into account under reproducibility conditions

7.4.3.3.1. Gauging error (or calibration error)

Whenever the gauging of an instrument (or the calibration of an absolute method) is not regularly redone, its output cannot be integrated in the reproducibility values. An experiment schedule must be carried out in order to estimate it using the residual error of the regression.

7.4.3.3.1.1. Procedure

The approach is similar to that carried out in the linearity study of the method.

It is recommended to implement a number ***n*** of reference materials. The number must be higher than 3, but it is not necessary to go beyond 10. The reference materials are to be measured ***p*** times under intralaboratory precision conditions, ***p*** must be higher than 3, a figure of 5 is generally recommended. The accepted values of reference materials must be regularly distributed on the range of values under study. The number of measurements must be the same for all the reference materials.

The results are reported in a table presented as follows:

| **Reference materials** | **Accepted value of the reference material** | **Measured values** | | | | |
|---|---|---|---|---|---|---|
| | | **Replica 1** | **...** | **Replica j** | **...** | **Replica p** |
| *1* | $x_1$ | $y_{11}$ | ... | *y1j* | ... | *y1p* |

| | | | | | | |
|---|---|---|---|---|---|---|
| ... | ... | ... | ... | ... | ... | ... |
| *i* | $x_i$ | $y_{i1}$ | ... | $y_{ij}$ | ... | *yip* |
| ... | ... | .... | ... | ... | ... | ... |
| *n* | *xn* | *yn1* | ... | *ynj* | ... | *ynp* |

7.4.3.3.1.2. Calculations and results

The linear regression model is calculated.

$$y_{ij} = a + b.x_i + \varepsilon_{ij}$$

Where

$y_{ij}$ is $j^{th}$ replica of the $i^{th}$ reference material.

$x_i$ is the accepted value of the $i^{th}$ reference material.

b is the slope of the regression line.

A is the intercept point of the regression line.

$a + b.x_i$ represent the expectation of the measurement value of the $\boldsymbol{i^{th}}$ reference material.

$\varepsilon_{ij}$ is the difference between $y_{ij}$ and the expectation of the measurement value of the $\boldsymbol{i^{th}}$ reference material.

The parameters of the regression line are obtained using the following formulae:

- mean of $\boldsymbol{p}$ measurements of the $\boldsymbol{i^{th}}$ reference material

$$yi = \frac{1}{p}\sum_{j=1}^{p} y_{ij}$$

- mean of all the accepted values of *n* reference materials

$$M_x = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- mean of all measurements

$$M_y = \frac{1}{n}\sum_{i=1}^{n} y_i$$

- estimated slope ***b***

$$b = \frac{\sum_{i-1}^{n}(x_i - M_x)(y_i - M_y)}{\sum_{i-1}^{n}(x_i - M_x)^2}$$

- estimated intercept point ***a***

$$a = M_y - b \times M_x$$

- regression value associated with the $i^{th}$ reference material $\hat{y}_i$

$$\hat{y}_i = a + b + x_i$$

- residual $e_{ij}$

$$e_{ij} = y_{ij} - \hat{y}_i$$

Estimating the standard uncertainty associated the gauging line (or calibration line)

If the errors due to the regression line are constant over the entire field, the standard uncertainty is estimated in a global, single way by the overall residual standard deviation.

$$u_{(gauging)} = S_{res} = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{p}(y_{ij} - \hat{y}_i)^2}{np - 2}}$$

If the errors due to the regression line are not constant over the entire field, the standard uncertainty is estimated for a given level by the residual standard deviation for this level.

$$u_{(gauging),i} = S_{res,i} = \sqrt{\frac{\sum_{j=1}^{p}(y_{ij} - \hat{y}_i)^2}{p - 1}}$$

Note These estimates of standard deviations can be used if the linear regression

model and the gauging (or calibration) domain have been validated (see § 5.3.1)

7.4.3.3.2. Bias error

According to the EURACHEM guide, *"Quantifying uncertainty in analytical measurements"*, it is recalled that the ISO guide generally requires that corrections be applied for all identified significant systematic effects. The same applies to the bias of methods for which the laboratory implements its quality control system (see §6), and which tends towards 0 for methods "under control".

In practice, a distinction can be made between two cases:

7.4.3.3.2.1. Methods adjusted with only one certified reference material

Bias is permanently adjusted with the same reference material.

The certified reference material (CRM) ensures the metrological traceability of the method. A reference value was allotted to the CRM together with its standard uncertainty $\boldsymbol{u_{ref}}$. This standard uncertainty of the CRM is combined with the compound uncertainty for the method, $u_{comp}$, to determine the overall standard uncertainty of the laboratory method ***u(x)***.

The overall standard uncertainty of the method adjusted with the CRM in question is therefore:

$$u(x) = \sqrt{u_{ref}^2 + u_{comp}^2}$$

Note 1 The methodology is identical in the case of methods adjusted with the results of an interlaboratory comparison chain.

Note 2 Note the difference between a CRM used to adjust the bias of a method, in which the uncertainty of its reference value combines with that of the method, and a CRM used to control a method adjusted by other means (cf. § 6.5.4.2). In the second case, the uncertainty of the CRM should not be used for the uncertainty assessment of the method.

7.4.3.3.2.2. Methods adjusted with several reference materials (gauging ranges etc.)

There is no particular adjustment of bias apart from gauging work.

It is clear that each calibrator introduces bias uncertainty. There is therefore an overall theoretical uncertainty of bias, which is a combination of the uncertainties of each calibrator. This uncertainty is very delicate to estimate, but it generally proves to be sufficiently low to be ignored, in particular if the laboratory monitors the quality of its calibrators, and the uncertainty of their reference values.

Other than in specific cases, bias uncertainty is ignored here.

7.4.3.3.3. Matrix effect

7.4.3.3.3.1. Definition

The matrix effect incurs a repeatable source of error for a given sample, but random from one sample to another. This error is related to the interaction of the compounds present in the product to be analyzed on measuring the required analyte. The matrix effect appears in methods with a nonspecific signal.

The matrix effect often constitutes a small part of uncertainty, particularly in separative methods. In certain other methods, including the infra-red techniques, it is a significant component of uncertainty.

**Example:** Estimate of the matrix effect on FTIR

The signal for the FTIR, or infra-red spectrum, is not a signal specific to each of the compounds that are measured by this technique. The statistical gauging model can be used to process disturbed, nonspecific spectral data in a sufficiently exact estimate of the value of the measurand. This model integrates the influences of the other compounds of the wine, which vary from one wine to the next and introduce an error into the result. Upstream of the routine analysis work, special work is carried out by the gauging developers to minimize this matrix effect and to make gauging robust, i.e. capable of integrating these variations without reflecting them in the result. Nevertheless the matrix effect is always present and constitutes a source of error at the origin of a significant part of the uncertainty of an FTIR method.

To be completely rigorous, this matrix effect error can be estimated by comparing, on the one hand, the means for a great number of FTIR measurement replicas, obtained on several reference materials (at least 10), under reproducibility conditions, and the true values of reference materials with a natural wine matrix on the other. The standard deviation of the differences gives this variability of gauging (provided that the gauging has been adjusted beforehand (bias = 0)).

This theoretical approach cannot be applied in practice, because the true values are never known, but it is experimentally possible to come sufficiently close to it:

- As a preliminary, the FTIR gauging must be statistically adjusted (bias = 0) in relation to a reference method based on at least 30 samples. This can be used to eliminate the effects of bias in the measurements thereafter.
- The reference materials must be natural wines. It is advisable to use at least 10 different reference materials, with values located inside a range level, the uncertainty of which can be considered to be constant.
- An acceptable reference value is acquired, based on the mean of several

measurements by the reference method, carried out under reproducibility conditions. This can be used to lower the uncertainty of the reference value: if, for the reference method used, all the significant sources of uncertainty range within reproducibility conditions, the multiplication of the number p of measurements carried out under reproducibility conditions, enable the uncertainty associated with their mean to be divided by $\sqrt{p}$. The mean obtained using a sufficient number of measurements will then have a low level of uncertainty, even negligible in relation to the uncertainty of the alternative method; and can therefore be used as a reference value. p must at least be equal to 5.

- The reference materials are analyzed by the FTIR method, with several replicas, acquired under reproducibility conditions. By multiplying the number of measurements q under reproducibility conditions on the FTIR method, the variability related to the precision of the method (random error) can be decreased. The mean value of these measurements will have a standard deviation of variability divided by $\sqrt{q}$. This random error can then become negligible in relation to the variability linked to the gauging (matrix effect) that we are trying to estimate. q must at least be equal to 5.

The following example is applied to the determination of acetic acid by FTIR gauging. The reference values are given by 5 measurements under reproducibility conditions on 7 stable test materials. The number of 7 materials is in theory insufficient, but the data here are only given by way of an example.

| | *Reference method* | | | | | | *FTIR* | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Materials | 1 | 2 | 3 | 4 | 5 | Mean Ref | 1 | 2 | 3 | 4 | 5 | Mean FTIR | Diff |
| *1* | *0.30* | *0.32* | *0.31* | *0.30* | *0.31* | *0.308* | *0.30* | *0.31* | *0.31* | *0.30* | *0.30* | *0.305* | ***-0.004*** |
| *2* | *0.31* | *0.32* | *0.32* | *0.32* | *0.31* | *0.316* | *0.31* | *0.32* | *0.30* | *0.31* | *0.31* | *0.315* | ***-0.006*** |
| *3* | *0.38* | *0.39* | *0.39* | *0.38* | *0.38* | *0.384* | *0.37* | *0.37* | *0.37* | *0.37* | *0.36* | *0.37* | ***-0.016*** |
| *4* | *0.25* | *0.25* | *0.25* | *0.24* | *0.25* | *0.248* | *0.26* | *0.26* | *0.26* | *0.25* | *0.26* | *0.26* | ***0.01*** |
| *5* | *0.39* | *0.39* | *0.40* | *0.40* | *0.39* | *0.394* | *0.43* | *0.42* | *0.43* | *0.42* | *0.42* | *0.425* | ***0.03*** |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *6* | *0.27* | *0.26* | *0.26* | *0.26* | *0.26* | *0.262* | *0.25* | *0.26* | *0.25* | *0.25* | *0.26* | *0.255* | ***-0.008*** |
| *7* | *0.37* | *0.37* | *0.37* | *0.37* | *0.36* | *0.368* | *0.37* | *0.36* | *0.36* | *0.35* | *0.36* | *0.365* | ***-0.008*** |

Calculation of the differences: ***diff = Mean FTIR – Mean ref.***
The mean of the differences $M_d$ ***= 0.000*** verifies (good adjustment of the FTIR compared with the reference method***)***

The standard deviation of the differences, $S_d$ ***= 0.015.*** It is this standard deviation that is used to estimate the variability generated by the gauging, and we can therefore state that:

$U_f$ ***= 0.015***

NOTE It should be noted that the value of $U_f$ can be over-estimated by this approach. If the laboratory considers that the value is significantly excessive under the operating conditions defined here, it can increase the number of measurements on the reference method and/or the FTIR method.
The reproducibility conditions include all the other significant sources of error, $S_R$ was otherwise calculated: ***SR = 0.017***

The uncertainty of the determination of acetic acid by this FTIR application is:

$$+\frac{}{-2} * \sqrt{0.015^2 + 0.017^2} + / - 0.045 g.L^{-1}$$

7.4.3.3.4. Sample effect

In certain cases, the experiment schedules used to estimate uncertainty are based on synthetic test materials. In such a situation, the estimate does not cover the sample effect (homogeneity). The laboratories must therefore estimate this effect.
It should be noted, however, that this effect is often negligible in oenological laboratories, which use homogeneous samples of small quantities.

7.4.4. Estimating standard uncertainty by interlaboratory tests

7.4.4.1. Principle

The interlaboratory approach uses data output by interlaboratory tests from which a

standard deviation of interlaboratory reproducibility is calculated, in accordance with the principles indicated in §5.4.3. The statisticians responsible for calculating the results of the interlaboratory tests can identify "aberrant" laboratory results, by using tests described in the ISO 5725 standard (Cochran test). These results can then be eliminated after agreement between the statisticians and the analysts.

For the uncertainty assessment by interlaboratory approach, the guidelines stated in the ISO 21748 standard are as follows:

- The reproducibility standard deviation (interlaboratory) obtained in a collaborative study is a valid basis for evaluating the uncertainty of measurement
- Effects that are not observed as part of the collaborative study must be obviously negligible or be explicitly taken into account.

There are two types of interlaboratory tests:

- Collaborative studies which relate to only one method. These studies are carried out for the initial validation of a new method in order to define the standard deviation of interlaboratory reproducibility $\boldsymbol{SR_{inter}}$ ***(method)***.
- Interlaboratory comparison chains, or aptitude tests. These tests are carried out to validate a method adopted by the laboratory, and the routine quality control (see § 5.3.3.3). The data are processed as a whole, and integrate all the analysis methods employed by the laboratories participating in the tests. The results are the interlaboratory mean $\boldsymbol{m}$, and the standard deviation of interlaboratory and intermethod reproducibility $\boldsymbol{SR_{inter}}$.

7.4.4.2. Using the standard deviation of interlaboratory and intramethod reproducibility $\boldsymbol{SR_{inter}}$ ***(method)***

The standard deviation of intralaboratory reproducibility $\boldsymbol{SR_{inter}}$ ***(method)*** takes into account intralaboratory variability and the overall interlaboratory variability related to the method.

Then must be taken into account the fact that the analysis method can produce a systematic bias compared with the true value.

As part of a collaborative study, whenever possible, the error produced by this bias can be estimated by using certified reference materials, under the same conditions as described in § 7.4.3.3.2, and added to $\boldsymbol{SR_{inter}}$ ***(method)***.

7.4.4.3. Using the standard deviation of interlaboratory and intermethod

reproducibility $SR_{inter}$

The standard deviation of intralaboratory reproducibility $SR_{inter}$ takes into account intralaboratory variability and interlaboratory variability for the parameter under study.

The laboratory must check its accuracy in relation to these results (see § 5.3.3).

There is no need to add components associated with method accuracy to the uncertainty budget, since in the "multi-method" aptitude tests, errors of accuracy can be considered to be taken into account in $SR_{inter}$.

7.4.4.4. Other components in the uncertainty budget

Insofar as the test materials used for the interlaboratory tests are representative of the conventional samples analyzed by laboratories, and that they follow the overall analytical procedure (sub-sampling, extraction, concentration, dilution, distillation etc.), $S_{R\text{-}inter}$ represents the standard uncertainty ***u(x)*** of the method, in the interlaboratory sense.

Errors not taken into account in the interlaboratory tests must then be studied in order to assess their compound standard uncertainty, which will be combined with the compound standard uncertainty of the interlaboratory tests.

7.5. Expressing expanded uncertainty

In practice, uncertainty is expressed in its expanded form, is absolute terms for methods in which uncertainty is stable in the scope in question, or relative when uncertainty varies proportionally in relation to the quantity of the measurand:

**Absolute uncertainty:** $U = +/- 2.u(x)$ $U = +/- 2.u(x)$

**Relative uncertainty (in %):** $U = + \frac{2.u(x)}{x}.100$

where $\bar{x}$ mean represents the reproducibility results.

Note This expression of uncertainty is possible given the assumption that the variations obey a normal law with a 95% confidence rate.

These expressions result in a given uncertainty value with a confidence level of **95%**.

**REFERENCES**


- **OIV, 2001** – Recueil des methods internationales d'analyse des vins and des moûts; *OIV Ed., Paris.*
- **(2) OIV, 2002** – Recommandations harmonisées pour le contrôle interne de qualité dans les laboratoires d'analyse; *OIV resolution œno 19/2002., Paris.*

- **(3) Standard ISO 5725: 1994** – Exactitude (justesse and fidélité) des results and methods de mesure, *classification index X 06-041-1*
- **(4) IUPAC, 2002** – Harmonized guidelines for single-laboratory validation of analysis methods; *Pure Appl. Chem., Vol. 74; n°5, pp. 835-855.*
- **(5) Standard ISO 11095: 1996** – Etalonnage linéaire utilisant des materials de référence, *reference number ISO 11095:1996*
- **(6) Standard ISO 21748: 2004** – Lignes directrices relatives à l'utilisation d'estimation de la répétabilité, de la reproductibilité and de la justesse dans l'évaluation de l'incertitude de mesure, *reference number ISO ISO/TS 21748:2004*
- **(7) Standard AFNOR V03-110: 1998** – Procédure de validation intralaboratory d'une method alternative par rapport à une method de référence, *classification index V03-110*
- **(8) Standard AFNOR V03-115: 1996** – Guide pour l'utilisation des materials de référence, *classification index V03-115*
- **(9) Standard AFNOR X 07-001: 1994** – Vocabulaire international des termes fondamentaux and généraux de métrologie, *classification index X07-001*
- **(10) Standard AFNOR ENV 13005: 1999** – Guide pour l'expression de l'incertitude de mesure
- **(11) AFNOR, 2003**, - Métrologie dans l'entreprise, outil de la qualité 2ème édition, *AFNOR 2003 édition*
- **(12) EURACHEM, 2000.** - Quantifying Uncertainty in Analytical Measurement, *EURACHEM second edition 2000*
- **(13) CITAC / EURACHEM, 2000** - Guide pour la qualité en chimie analytique, *EURACHEM 2002 edition*
- **(14) Bouvier J.C., 2002** - Calcul de l'incertitude de mesure – Guide pratique pour les laboratoires d'analyse œnologique, *Revue Française d'œnologie no.197, Nov-Dec 2002, pp: 16-21*
- **(15) Snakkers G. and Cantagrel R., 2004** - Utilisation des données des circuits de comparaison interlaboratoires pour apprécier l'exactitude des results d'un laboratoire Estimation d'une incertitude de mesure - *Bull OIV, Vol. 77 857-876, Jan – Feb 2004, pp: 48-83*

- **(16) Perruchet C. and Priel M, 2000** - Estimer l'incertitude, *AFNOR Editions*
- **(17) Neuilly (M.) and CETAMA, 1993** - Modélisation and estimation des errors de mesures, *Lavoisier Ed, Paris*


**Annex N°1**

**Table A -Law of SNEDECOR**

This table indicates values of **F** in function with **□1** and **□2** for a risk **□□**of 0,05

P=0,950

| **□□1 □2** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **□1 □□□2** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 161,4 | 199,5 | 215,7 | 224,6 | 230,2 | 234,0 | 236,8 | 238,9 | 240,5 | 241,9 | **1** |
| **2** | 18,51 | 19,00 | 19,16 | 19,25 | 19,30 | 19,33 | 19,35 | 19,37 | 19,38 | 19,40 | **2** |
| **3** | 10,13 | 9,55 | 9,28 | 9,12 | 9,01 | 8,94 | 8,89 | 8,85 | 8,81 | 8,79 | **3** |
| **4** | 7,71 | 6,94 | 6,59 | 6,39 | 6,26 | 6,16 | 6,09 | 6,04 | 6,00 | 5,96 | **4** |
| **5** | 6,61 | 5,79 | 5,41 | 5,19 | 5,05 | 4,95 | 4,88 | 4,82 | 4,77 | 4,74 | **5** |
| **6** | 5,99 | 5,14 | 4,76 | 4,53 | 4,39 | 4,28 | 4,21 | 4,15 | 4,10 | 4,06 | **6** |
| **7** | 5,59 | 4,74 | 4,35 | 4,12 | 3,97 | 3,87 | 3,79 | 3,73 | 3,68 | 3,64 | **7** |
| **8** | 5,32 | 4,46 | 4,07 | 3,84 | 3,69 | 3,58 | 3,50 | 3,44 | 3,39 | 3,35 | **8** |
| **9** | 5,12 | 4,26 | 3,86 | 3,63 | 3,48 | 3,37 | 3,29 | 3,23 | 3,18 | 3,14 | **9** |
| **10** | 4,96 | 4,10 | 3,71 | 3,48 | 3,33 | 3,22 | 3,14 | 3,07 | 3,02 | 2,98 | **10** |
| **11** | 4,84 | 3,98 | 3,59 | 3,36 | 3,20 | 3,09 | 3,01 | 2,95 | 2,90 | 2,85 | **11** |
| **12** | 4,75 | 3,89 | 3,49 | 3,26 | 3,11 | 3,00 | 2,91 | 2,85 | 2,80 | 2,75 | **12** |
| **13** | 4,67 | 3,81 | 3,41 | 3,18 | 3,03 | 2,92 | 2,83 | 2,77 | 2,71 | 2,67 | **13** |
| **14** | 4,60 | 3,74 | 3,34 | 3,11 | 2,96 | 2,85 | 2,76 | 2,70 | 2,65 | 2,60 | **14** |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **15** | 4,54 | 3,68 | 3,29 | 3,06 | 2,90 | 2,79 | 2,71 | 2,64 | 2,59 | 2,54 | **15** |
| **16** | 4,49 | 3,63 | 3,24 | 3,01 | 2,85 | 2,74 | 2,66 | 2,59 | 2,54 | 2,49 | **16** |
| **17** | 4,45 | 3,59 | 3,20 | 2,96 | 2,81 | 2,70 | 2,61 | 2,55 | 2,49 | 2,45 | **17** |
| **18** | 4,41 | 3,55 | 3,16 | 2,93 | 2,77 | 2,66 | 2,58 | 2,51 | 2,46 | 2,41 | **18** |
| **19** | 4,38 | 3,52 | 3,13 | 2,90 | 2,74 | 2,63 | 2,54 | 2,48 | 2,42 | 2,38 | **19** |
| **20** | 4,35 | 3,49 | 3,10 | 2,87 | 2,71 | 2,60 | 2,51 | 2,45 | 2,39 | 2,35 | **20** |
| **21** | 4,32 | 3,47 | 3,07 | 2,84 | 2,68 | 2,57 | 2,49 | 2,42 | 2,37 | 2,32 | **21** |
| **22** | 4,30 | 3,44 | 3,05 | 2,82 | 2,66 | 2,55 | 2,46 | 2,40 | 2,34 | 2,30 | **22** |
| **23** | 4,28 | 3,42 | 3,03 | 2,80 | 2,64 | 2,53 | 2,44 | 2,37 | 2,32 | 2,27 | **23** |
| **24** | 4,26 | 3,40 | 3,01 | 2,78 | 2,62 | 2,51 | 2,42 | 2,36 | 2,30 | 2,25 | **24** |
| **25** | 4,24 | 3,39 | 2,99 | 2,76 | 2,60 | 2,49 | 2,40 | 2,34 | 2,28 | 2,24 | **25** |
| **26** | 4,23 | 3,37 | 2,98 | 2,74 | 2,59 | 2,47 | 2,39 | 2,32 | 2,27 | 2,22 | **26** |
| **27** | 4,21 | 3,35 | 2,96 | 2,73 | 2,57 | 2,46 | 2,37 | 2,31 | 2,25 | 2,20 | **27** |
| **28** | 4,20 | 3,34 | 2,95 | 2,71 | 2,56 | 2,45 | 2,36 | 2,29 | 2,24 | 2,19 | **28** |
| **29** | 4,18 | 3,33 | 2,93 | 2,70 | 2,55 | 2,43 | 2,35 | 2,28 | 2,22 | 2,18 | **29** |
| **30** | 4,17 | 3,32 | 2,92 | 2,69 | 2,53 | 2,42 | 2,33 | 2,27 | 2,21 | 2,16 | **30** |
| **40** | 4,08 | 3,23 | 2,84 | 2,61 | 2,45 | 2,34 | 2,25 | 2,18 | 2,12 | 2,08 | **40** |
| **60** | 4,00 | 3,15 | 2,76 | 2,53 | 2,37 | 2,25 | 2,17 | 2,10 | 2,04 | 1,99 | **60** |
| **120** | 3,92 | 3,07 | 2,68 | 2,45 | 2,29 | 2,17 | 2,09 | 2,02 | 1,96 | 1,91 | **120** |
| **∞** | 3,84 | 3,00 | 2,60 | 2,37 | 2,21 | 2,10 | 2,01 | 1,94 | 1,88 | 1,83 | **∞** |

| □2<br>□□□1 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | □□2<br>□1 |
|---|---|---|---|---|---|---|---|---|---|---|---|